Title
Improved methods for single-particle cryogenic electron microscopy

Author
Palovcak, Eugene Joseph

Publication Date
2019

Peer reviewed|Thesis/dissertation

Improved methods for single-particle cryogenic electron microscopy

by
Eugene Joseph Palovcak

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Yifan Cheng
37327FF78C17460...
Yifan Cheng
Chair

DocuSigned by:
Geeta Narlikar
Geeta Narlikar

DocuSigned by:
David A. Agard
328308DD7E5546A...
David A. Agard

Committee Members

Acknowledgements

I would like to acknowledge some of the people I've had the pleasure of knowing here:

Vincenzo Carnevale, my first teacher and mentor. I hope I can pay forward his generosity.

My friends from Pennsylvania, who cheered me on when I jumped across the country to live in California alone and do the hardest thing I've ever done.

James Fraser, who is a model of what a scientist and mentor and good person should be.

Yifan Cheng, who has a beginner's mind and a runner's resolve. I aspire to his good nature, his perpetual enthusiasm for science, and his kindness.

Evan Green, who has convinced me that time worrying is better spent with friends outside.

James Palovcak, who taught me about duty and love by example.

Lillian Kenner, my partner. Without you, I would be unrecognizable to myself.
You fill my life with light and make me better.

Ronin Palovcak, it's all for you now, kid.

Contributions

Chapter 1

Zheng, S.Q., **Palovcak, E.**, Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods 14*, 331-332.

Chapter 2

Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A., Cheng, Y. (2018). A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *Journal of Structural Biology.* 204(1), 80-84.

Chapter 3

Palovcak, E., Asarnow, D., Campbell, M., Yu, Z., Cheng, Y. (2019). Generating contrast in cryo-EM images with convolutional networks. *In submission.*

Epigraph

Everything is made of atoms. That is the key hypothesis. The most important hypothesis in all of biology, for example, is that everything that animals do, atoms do. In other words, there is nothing that living things do that cannot be understood from the point of view that they are made of atoms acting according to the laws of physics.
-- Richard Feynman, Atoms in Motion

"Faith" is a fine invention
For Gentlemen who *see*!
But Microscopes are prudent
In an Emergency!
-- Emily Dickinson

If I had nine hours to chop down a tree, I'd spend the first six sharpening my axe.
-- mythically attributed to Abraham Lincoln

Improved methods for single-particle cryogenic electron microscopy

Eugene Joseph Palovcak

Abstract

Biological macromolecules such as enzymes are nanoscale machines. This is true in a concrete sense: if the atomic structure of a biological macromolecule can be obtained, the theories of mechanics and intermolecular forces can be applied to explain how the machine works in terms that engineers would understand, including motors, ratchets, gates and transducers. Nevertheless, biological macromolecules are complex, fragile and extremely small, so obtaining their structures is a challenging experimental endeavor.

Single-particle cryogenic electron microscopy (cryo-EM) is a technique for determining the 3D structure of a biological macromolecule from a large set of 2D electron micrographs of individual structurally-identical particles. To obtain such images, a solution of the macromolecules must be prepared in the frozen-hydrated state, embedded in a thin electron-transparent glassy film of water. This specimen must then be imaged with a very short exposure to avoid radiation damage. A powerful computer must then be used to sort, align, and average the 2D particle images to back-calculate the 3D structure. At its best, cryo-EM can determine the structures of biological macromolecules to atomic resolution. In practice, this goal is usually not achieved.

Cryo-EM has gotten significantly more powerful in the past few years due to improvements in equipment and methodology. Several of the most significant advances originated in the labs of David Agard and Yifan Cheng at UCSF. When I began my PhD with Yifan, the spirit in the lab was that cryo-EM could keep getting better and better: with enough engineering, determining the

3D structure of an arbitrary biological macromolecule would be as routine an experiment as gel electrophoresis or DNA sequencing.

Inspired, I took on projects in the lab that I thought would move the field closer to that goal. In the first chapter of this thesis, I describe work I did supporting a project initiated by David Agard and his long-time scientific programmer Shawn Zheng. They developed and implemented an algorithm, MotionCor2, for correcting the complex, anisotropic movements that occur when a frozen-hydrated specimen interacts with the high-energy electron beam. My role was to benchmark MotionCor2 on a panel of real-world 3D reconstruction tasks. I was able to show that MotionCor2 restored the highest resolution details in the images, ultimately yielding significantly better structures than simpler algorithms. For me, this projected highlighted the importance of benchmarking an algorithm for use in routine real-world conditions with the right metrics. In chapter 1, I include the manuscript for the MotionCor2 study, formatted to highlight my contributions that were moved to the supplement in the original publication by Nature Methods.

One of the major remaining issues with cryo-EM is sample preparation: preparing the thin freestanding films of frozen-hydrated particles necessarily exposes those particles to air-water interfaces. Many fragile macromolecular complexes denature when exposed to such interfaces, preventing structure determination with cryo-EM. In chapters 2 and 3, I describe my efforts to develop a simple, robust approach to stabilizing fragile macromolecular complexes during the vitrification process. In chapter 2, I develop a method for coating EM grids with an electron-transparent and functionalizable graphene-oxide support film. I demonstrate that such GO grids are compatible with high-resolution structure determination. This work was published in the Journal of Structural Biology in 2018. In chapter 3, I extend this work by functionalizing GO grids with nucleic acids, enabling routine structure determination of uncrosslinked chromatin

specimens. In on-going work, I used nucleic acid grids to solve high-resolution structures of a highly fragile specimen, the snf2h-nucleosome complex, and analyzed the conformational heterogeneity of the nucleosome substrate. These results were made possible by the nucleic acid grid, as the other major approach for stabilizing chromatin specimens, chemical crosslinking, not work for this specimen.

Perhaps the most fundamental problem with single-particle cryo-EM is the radiation sensitivity of frozen-hydrated macromolecules. To image biological matter with electrons is to destroy it, so obtaining images of undamaged specimens requires very short, highly under sampled exposures. The resultant images are extremely noisy and low contrast, with most particles barely visible from the background. In chapter 4, I describe a novel computational approach to generating contrast in cryo-EM. Using a recently described machine learning strategy for training a parameterized denoising algorithm, I developed a computer program, *restore*, that denoises cryo-EM images, greatly enhancing their contrast and interpretability. This program leverages recent advances in computer vision and deep learning which have not yet been widely used in cryo-EM image processing algorithms. To characterize the performance of the algorithm on real-world data, I extended conventional metrics for image resolution to measure how an arbitrary transformation affects images at different spatial frequencies. These novel metrics are general and may be useful for characterizing other nonlinear reconstruction algorithms in cryo-EM and medical imaging. Finally, I showed that denoised cryo-EM images maintain the high-resolution information required for accurate 3D reconstruction. Denoising can be applied to conventional cryo-EM images and can be reversed whenever necessary. I have made the software for *restore* program publicly available and have submitted a manuscript for peer-reviewed publication.

Table of Contents

List of Figures

Chapter 1 - Motioncor2: Anisotropic correction of beam-induced motion

Contributing authors

Shawn Q. Zheng[1,2], Eugene Palovcak[1] , Jean-Paul Armache[1], Kliment Verba[1],

Yifan Cheng[1,2] and David A. Agard[1,2]

[1]Department of Biochemistry and Biophysics and [2]Howard Hughes Medical Institute, University of California, San Francisco, CA 94158

Introduction

Sample illumination with the high-energy electron beam breaks bonds, releases radiolysis products and builds up charge within the thin frozen hydrated biological samples during image recording. The result is a combination of physical and optical distortions that can significantly deteriorate sample high-resolution information through image blurring. This has been one of the major factors limiting the achievable resolution of single particle cryo-EM. The concept of recording the image as movie to correct sample motion was proposed long ago (Henderson and Glaeser, 1985; Typke et al., 2004), but only became practical once direct electron detection cameras became available (Campbell et al., 2012). Motion can be measured by tracking the movement of images captured in a series of snapshots, either as the whole frame or as individual particles. Motion can then be corrected by registering identical features in the sub-frames to each other, followed by summing the registered sub-frames to produce a motion-corrected image. While in principle tracking should be straightforward, the practical challenge is the extremely low signal-

to-noise ratio (SNR) in each individual sub-frame. Except for very large particles, accurate motion measurement generally requires correlating the motion between sub-frames over large areas. That said, even sub-optimal motion correction can significantly restore high-resolution signals and improve the resolution of final 3D reconstructions (Campbell et al., 2012; Li et al., 2013).

We previously developed an algorithm to correct uniform image motion (Li et al., 2013). It uses redundant measurements of image shifts between all sub-frames to derive a least-squares estimate of relative motions between neighboring sub-frames. We implemented this algorithm in the program MotionCorr that provided an efficient correction of image motions with sufficient accuracy to enable the determination of numerous near atomic resolution 3D reconstructions (Liao et al., 2013, Paulsen et al., 2015). Around the same time, a number of different strategies were devised that either assume particles located nearby have similar motions or assume uniform motion of the entire frame or patches of the frame. Programs based on the former assumption include RELION that provides a movie-processing mode but uses a 3D reconstruction to track particle motions (Scheres, 2012), Xmipp that implemented an Optic Flow algorithm (Abrishami et al., 2015), and alignparts_lmbfgs that implemented a regularized Fourier Space optimization algorithm to track neighboring particles (Rubinstein and Brubaker, 2015). Programs based on tracking the full frame or parts of the frame include MotionCorr and iterative whole frame alignment procedures such as Unblur (Li et al., 2013; Grant and Grigorieff, 2015). Each of these algorithms have demonstrated the ability to recover high-resolution signals to varying degrees and have improved the resolution of the resultant 3D reconstructions.

Ideally, single particle cryo-EM images should be acquired with the smallest possible defocus to enhance high-resolution information, and with the shortest possible sub-frame exposure times to reduce motion trapped within individual sub-frames, in particular for the first few frames

where the sample has the least radiation damage but moves most rapidly (Wu et al., 2015). Additionally, motion detection should be done on the smallest possible local area to best capture the anisotropic motion. Unfortunately, minimizing defocus, time or area significantly reduces the SNR available for measurement, ultimately leading to incorrect motion estimates. For example, previous experiments with MotionCorr revealed that subdividing the images into areas smaller than ~2000 × ~2000 pixels or going to sub-frame integration times of less than 100 milliseconds worsened resolution due to increased errors in motion tracking.

In this work we experimentally validated the doming behavior that was proposed to describe beam-induced sample motion and discovered that the initial motion begins at surprisingly low doses (Brilot et al., 2012). Our observations support a model in which the sample is smoothly deformed throughout the exposure. The extreme sensitivity to dose emphasizes the need to acquire very short single frame exposures, further raising the demands for performance at low SNRs.

To meet the compounding demands for locality and low SNR, it was not possible to accurately determine local shifts by comparing noisy subregions of each frame to one another as done within MotionCorr. Instead, we compare local image regions to an approximated sum for each image subregion, and then iterate translation determination to improve the sum. The resultant shifts are further constrained both spatially and temporally, informed by the physical behavior of the sample. This is accomplished by deriving the final shifts from a fit of a time-varying two-dimensional (2D) polynomial function to the local motions derived from different patches of the image. Each image sub-frame is subsequently remapped using this smooth distribution of beam induced motions at each individual pixel and summed with or without radiation damage weighting. This algorithm has been implemented as MotionCor2, a parallel computing program running on Linux platform equipped with multiple GPUs. Our tests have shown that this program is very

robust, and sufficiently accurate at correcting local motions so that the very time-consuming and computationally intensive particle polishing procedure in RELION can be skipped.

Modeling of beam-induced motion

Early studies suggested that beam induced motion can be described as specimen doming (Brilot et al., 2012). This model predicts a significant sample motion perpendicular to the sample plane (z-motion, **Figure 1.1**), which, when the sample is tilted, should be projected onto the image plane as a motion perpendicular to the tilt axis (**Figure S1.1**). We explored this by directly measuring motions within movie stacks recorded from specimens tilted at various angles. We collected a number of dose fractionated tomographic tilt series from various frozen hydrated biological samples. By using rather low magnifications (4.3Å/pixel), we could accurately determine full frame shifts with doses down to 0.05 e- /Å2 using MotionCorr (Li et al., 2013). Representative traces of uniform whole frame motion of the same specimen at three tilt angles a=-30º, 0º, and 60º clearly show that the measured motion is much larger at higher tilt angles and mainly perpendicular to the tilt axis (**Figure S1.1**), supporting the dome model. Consistent with an earlier study[3], the motion within the very first 0.2 e-/A^2 is substantial. The beam-induced shifts observed during normal single particle cryo-EM must be a combination of sample drift, residual tilt, and non-idealities in the doming such that local motions are not purely along Z.

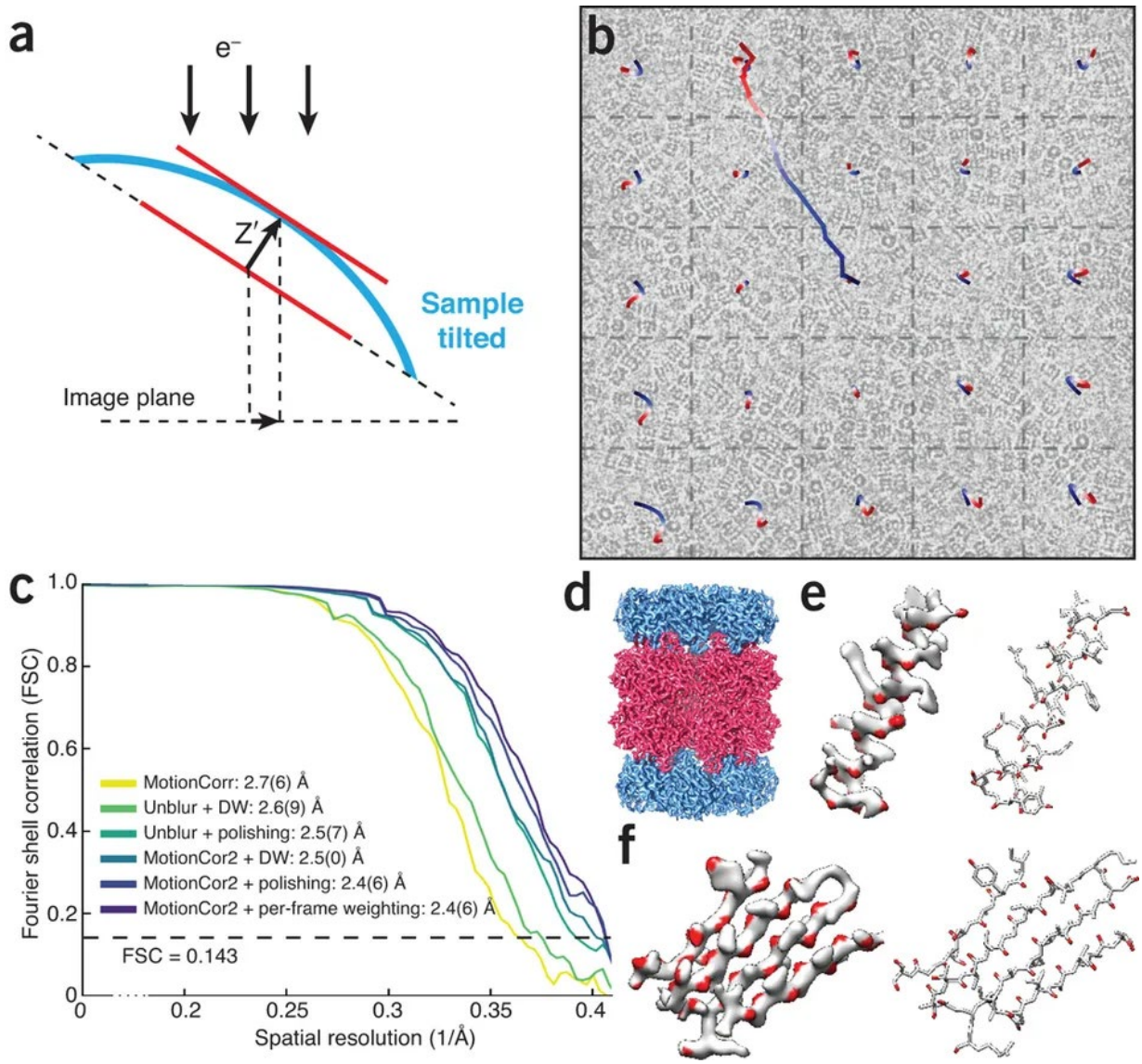

Figure 1.1: Local motion correction by MotionCor2. (a) Schematic drawing illustrates that when the sample is tilted the observable motion in the image plan is the projection of *z*-motion produced by doming of the sample under electron beam. (b) Image of frozen hydrated archaeal 20S proteasome overlaid with the traces of global motion based upon whole-frame alignment (long trace originated from the center of image) and each patch predicted from the polynomial function. (c) Fourier shell correlation (FSC) curves of 3D reconstructions determined using micrographs corrected by Unblur with dose weighting, Unblur followed by particle polishing, correction by MotionCor2 with dose weighting, MotionCor2 followed by particle polishing and MotionCor2 with per-frame B-factor weighting. (d) 3D reconstruction of archaeal 20S proteasome filtered to 2.5-Å resolution and sharpened by a temperature factor of −103.8 Å2. (e) Density of an α helix from the map, with resolved oxygen atom functional groups colored in red. Visualization of main chain carbonyls requires resolution below 3 Å. The refined atomic model is shown side by side for comparison. (f) As in e, but showing a β sheet.

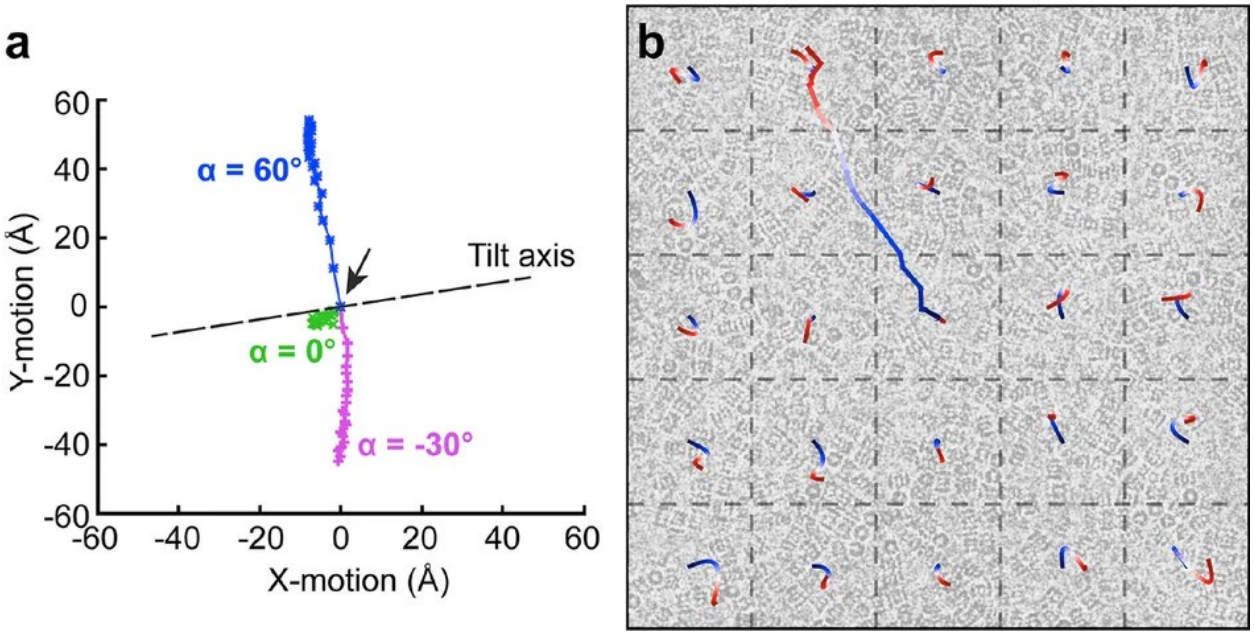

Figure S1.1: A doming model describes beam-induced motion (a) Traces of the projected motion measured at three different tilt angles (α-angle) extracted from a dose fractionated tilt series acquired on a frozen hydrated specimen of a wild-type Drosophila melanogaster γ-Tubulin Ring Complex. The black arrow in the plot indicates the starting position of the motion. (b) Image of frozen hydrated archaeal 20S proteasome overlaid with the traces of global motion based upon whole frame alignment (long trace originated from the center of image) and each patch determined from MotionCor2. The whole frame is divided into 5 × 5 patches, and traces of each patch are determined individually. The traces of the global and local motions are exaggerated on the image by a factor of 145. For perspective, the accumulated global motion is ~11 A.

Motion correction using polynomial constraints

Whole frame-based motion correction algorithms correct only the uniform motion across the entire image frame, leaving the residual non-uniform local motion uncorrected. A common strategy for correcting local motions is to spatially partition a movie stack into multiple stacks of patches where the correction is performed individually. Ideally, the smaller the patch size, the better the locality of motion correction. In practice, there is a limit on patch size, below which the motion cannot be accurately corrected due to insufficient SNR. Furthermore, dividing the whole image into a limited number of patches and correcting motion within each patch independently causes edge artifacts at the boundaries of patches in the corrected image, the so-called

'checkerboard' artifact. This creates additional challenges in single particle work and prevents use in tomography.

To avoid these problems, we chose to combine iterative patch-based motion measurement with restraints derived from the physical behavior of a smoothly doming surface to obtain a function that describes the motion of each pixel across the entire image. Since a dome can be geometrically approximated by a quadratic surface, we chose to fit the locally measured x and y shifts from all patches independently to a time-varying polynomial function that is quadratic in the xy plane and cubic with respect to exposure time t:

$$\begin{aligned} S(x,y,t) = {} & c_0 t + c_1 t^2 + c_2 t^3 + c_3 xt + c_4 xt^2 + c_5 xt^3 + c_6 x^2 t + c_7 x^2 t^2 + c_8 x^2 t^3 \\ & + c_9 yt + c_{10} yt^2 + c_{11} yt^3 + c_{12} y^2 t + c_{13} y^2 t^2 + c_{14} y^2 t^3 \\ & + c_{15} xyt + c_{16} xyt^2 + c_{17} xyt^3 \end{aligned} \qquad (1)$$

where, S is the shift at the integer pixel coordinate x and y. At any given time t, the above equation becomes a standard quadratic function:

$$S(x,y)_{t=\tau} = a_0 + a_a x + a_2 x^2 + a_3 y + a_4 y^2 + a_5 xy \qquad (2)$$

that approximates a domed surface. Since Eq. (1) is high-order differentiable over x, y and t, a smooth estimate of the shift can be obtained at each pixel of each sub-frame and between consecutive sub-frames. Not only does this strategy eliminate checkerboard artifacts, but importantly it also provides a more thorough correction of image motion. It effectively dampens the fluctuations of local measurements influenced by image noise.

Motion correction algorithm

Image motion can be decomposed into two components, the uniform global motion and the non-uniform local motion that is described as the projection of the doming motion. The motion correction is divided into two steps, global motion correction followed by local motion correction.

The global motion is corrected iteratively by calculating cross correlation of each subframe against the sum of all other sub-frames based upon the alignment of the previous iteration. To avoid self-correlation, the sub-frame being measured is excluded from the sum. The measured shifts, i.e., the residual errors of the previous iteration of alignment, are corrected by shifting the phase in the Fourier transforms of sub-frames. The iteration stops when the residual errors are below a specified tolerance (typically, 0.5 super-resolution pixel) or a specified maximum number of iterations have been reached.

The global-motion-corrected stack is partitioned into non-overlapping patches. The same alignment procedure is applied to each stack of patches. The local motion is therefore measured at a series of discrete spatial locations represented by the patch centers. The measured local shifts of each patch for each subframe are fit to a polynomial function given in Eq. (1). Least squares fitting is used to determine the unknown coefficients c_i. The measured x and y shifts are fit separately to Eq. (1), resulting two polynomial functions, $S_x(x,y,t)$ and $S_y(x,y,t)$, corresponding respectively to the x and y components of local motion. The vector function (S_x, S_y) provides a complete and smooth description of the local motion field, allowing smooth correction of the local motion at the single pixel level without causing edge effects.

Since the calculated shifts are non-integer and varying from pixel to pixel, the correction of pixel shifts must be performed in real space and we used bilinear interpolation. To minimize the attenuation of high-resolution signals due to interpolation, the interpolation is performed on

super-resolution pixels. The final image is obtained by cropping the corrected sum in Fourier domain to the user specified resolution.

Motion correction program MotionCor2

We first correct the uniform global motion by iteratively aligning each sub-frame against a reference that is the motion corrected sum of all other sub-frames based upon the translational alignment obtained in the previous cycle. Excluding the sub-frame being aligned from the reference prevents a strong auto-correlation peak at the origin that may influence the determination of the real cross correlation peak. The measured sub-frame shifts at each iteration are the residual errors from the last iteration, and the alignment procedure is terminated when the maximum residual error is below a specified tolerance. The measured global shifts are then corrected by phase shifting in the Fourier domain to yield a global-motion-corrected stack. The image stack is then divided into a grid of patches, and the same iterative alignment procedure is performed on each patch. This provides the local motion at a series of discrete locations for each time point within the exposure. These local motions are then fit to polynomial functions defined in Eq. (1). As a result, x and y shifts can be calculated according to the fit functions at each pixel (m, n) in each sub-frame. Since the calculated shifts at individual pixels are typically non-integer, the corresponding correction requires interpolation in real space. To minimize the attenuation of high-resolution signals due to interpolation, bilinear interpolation is performed on super-resolution pixels. The final image is obtained by cropping the corrected sum in Fourier domain to the user specified resolution. In practice the whole stack is typically divided into 5′5 partitions, an empirical choice that provides a good balance between precision, efficiency, and SNR. However, for the higher contrast Volta phase plate images, good results have been obtained using 9x9 patches

(Danev et al., 2016). In order to maximize the throughput of motion correction, parallel computation was implemented in three levels in MotionCor2 on a Linux platform equipped with multiple GPUs.

MotionCor2 improves accuracy of motion correction

We tested the performance of MotionCor2 using two previously acquired single particle cryo-EM datasets: the archaeal 20S proteasome and the rat TRPV1 ion channel. Both were previously processed using whole frame based MotionCorr and produced near atomic resolution structures (Li et al., 2013; Liao et al., 2013). These two datasets were collected using the same microscope settings and similar electron dose rates on the camera. As these were older data sets, the total dose was limited: the 20S stacks contained 25 sub-frames (total dose of ~34 electrons per $Å^2$) whereas the TRPV1 stacks had 30 sub-frames (total dose of ~41 electrons per $Å^2$). We re-processed motion correction for these two datasets using MotionCor2, configured to run on 5x5 patches.

MotionCor2 significantly improves Thon ring signals at high-resolution and leads to a better correlation with the simulated contrast transfer function (CTF) (**Figure S1.2, Figure S1.3**). Over the entire data set, almost all images showed a significant improvement in recovery of high resolution Thon ring signals. Performing a similar analysis with Unblur, a popular current-generation iterative whole frame motion correction strategy, shows that MotionCor2 provides a further improvement in Thon ring resolution (**Figure S1.2**).

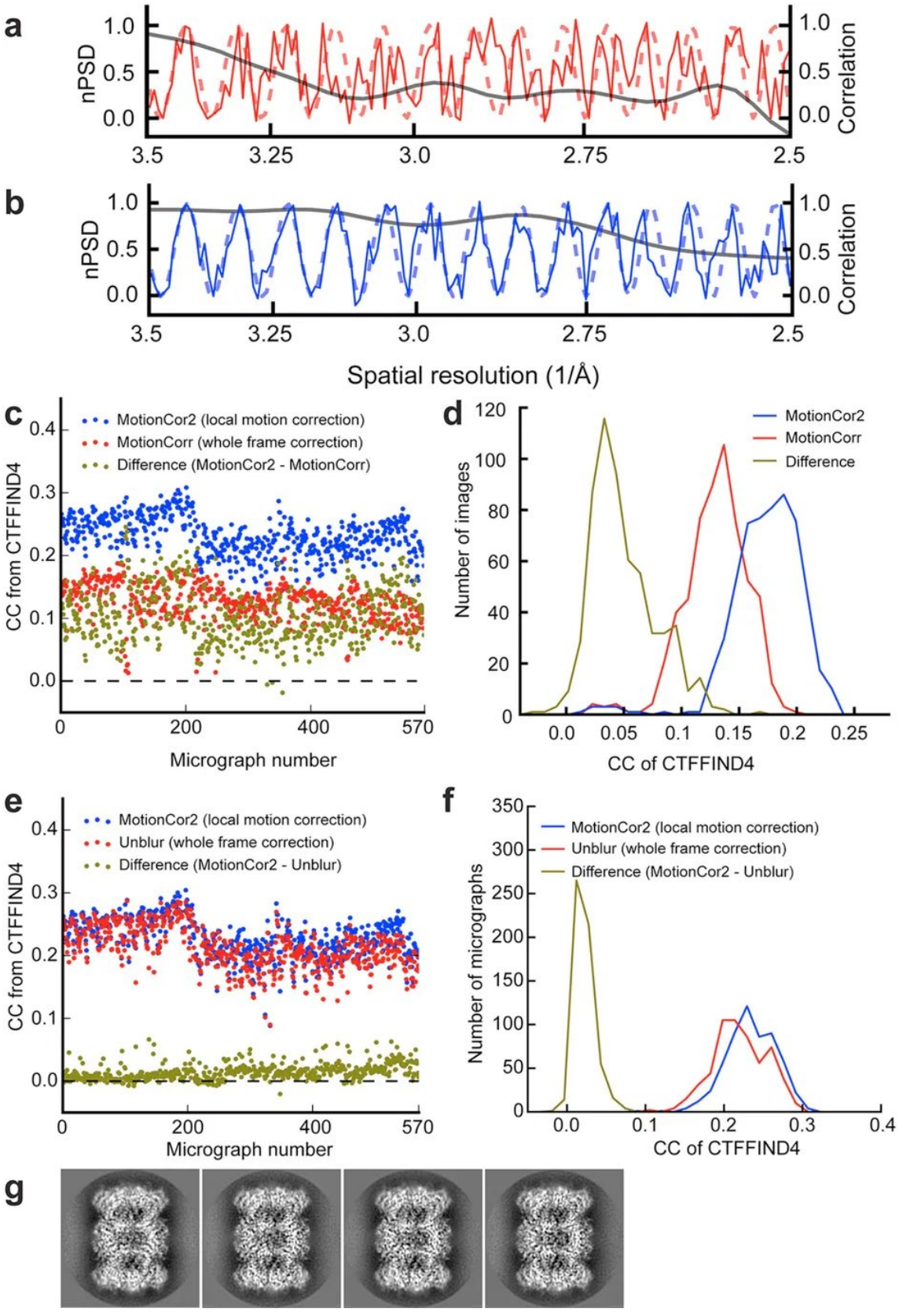
a
nPSD
Correlation
1.0
0.5
0.0
3.5
3.25
3.0
2.75
2.5
b
Spatial resolution (1/Å)
c
MotionCor2 (local motion correction)
MotionCorr (whole frame correction)
Difference (MotionCor2 - MotionCorr)
CC from CTFFIND4
Micrograph number
0
200
400
570
d
MotionCor2
MotionCorr
Difference
Number of images
CC of CTFFIND4
e
MotionCor2 (local motion correction)
Unblur (whole frame correction)
Difference (MotionCor2 - Unblur)
f
Number of micrographs
g

Figure S1.2: Comparison of motion correction algorithms on 20S proteasome. (a) Rotationally averaged Fourier power spectrum of image after motion correction by MotionCorr (solid line) and fitted contrast transfer function (dashed line). (b) Rotationally averaged Fourier power spectrum of image after motion correction by MotionCor2 (solid line) and fitted CTF (dashed line). Solid black line in both (a) and (b) indicate cross correlation coefficient between the rotationally averaged Fourier power spectrum of image and fitted CTF. (c) The cross correlation coefficient output from CTFFIND4 was used to assess the Thon ring quality within the resolution range of 10 ~ 5 Å. Blue and red dots represent cross correlation coefficient of every micrograph after motion correction using MotionCor2 (blue) and MotionCorr (red). Brown dots represent the differences between the two. (d) Histogram of cross correlation coefficients between calculated and simulated Fourier power spectrum of MotionCorr corrected image (red) and MotionCor2 corrected image (blue). The difference, which shows the amount of improvement, is shown in brown. (e) Comparison of motion corrections by MotionCor2 and Unblur. The cross-correlation coefficient output from CTFFIND4 calculated within the resolution range of 10 ~ 5 Å. Blue and red dots represent cross correlation coefficient of every micrograph after motion correction using MotionCor2 (blue) and Unblur (red). Brown dots represent the differences between the two. (f) Histogram of cross correlation coefficients between calculated and simulated Fourier power spectrum of MotionCor2 corrected image (blue) and Unblur corrected image (red). The difference, which shows the amount of improvement, is shown in brown. (g) 2D class averages of 20S proteasome calculated from images after motion correction by MotionCor2.

To better understand the impact of the improved motion correction, we re-determined the 3D reconstructions of the archaeal 20S proteasome from images that were processed by MotionCorr, Unblur and MotionCor2, followed by refinement and reconstruction with the maximum-likelihood based refinement and reconstruction program, RELION (**Figure 1.1**).

In the 3D reconstruction of T20S after MotionCor2 and dose weighting (**Figure 1.1**) most backbone carbonyls are now clearly visible, as well as the precise rotameric states of amino acid side chains (**Figure 1.1**). Such fine structural details facilitate more accurate model building.

Similarly, we re-processed our previous raw micrographs from the TRPV1 ion channel and re-determined its 3D reconstruction. MotionCor2 improves the nominal resolution from 3.5Å to 3.1(5)Å (**Figure S1.3**). Such improvements are particularly obvious in some trans-membrane regions, where extra densities associated with the TRPV1 protein are now seen to have well-defined features that can only now be interpreted as lipid molecules (**Figure S1.3**). Thus, for both the 20S proteasome and TRPV1 datasets, MotionCor2 produced noticeable improvements.

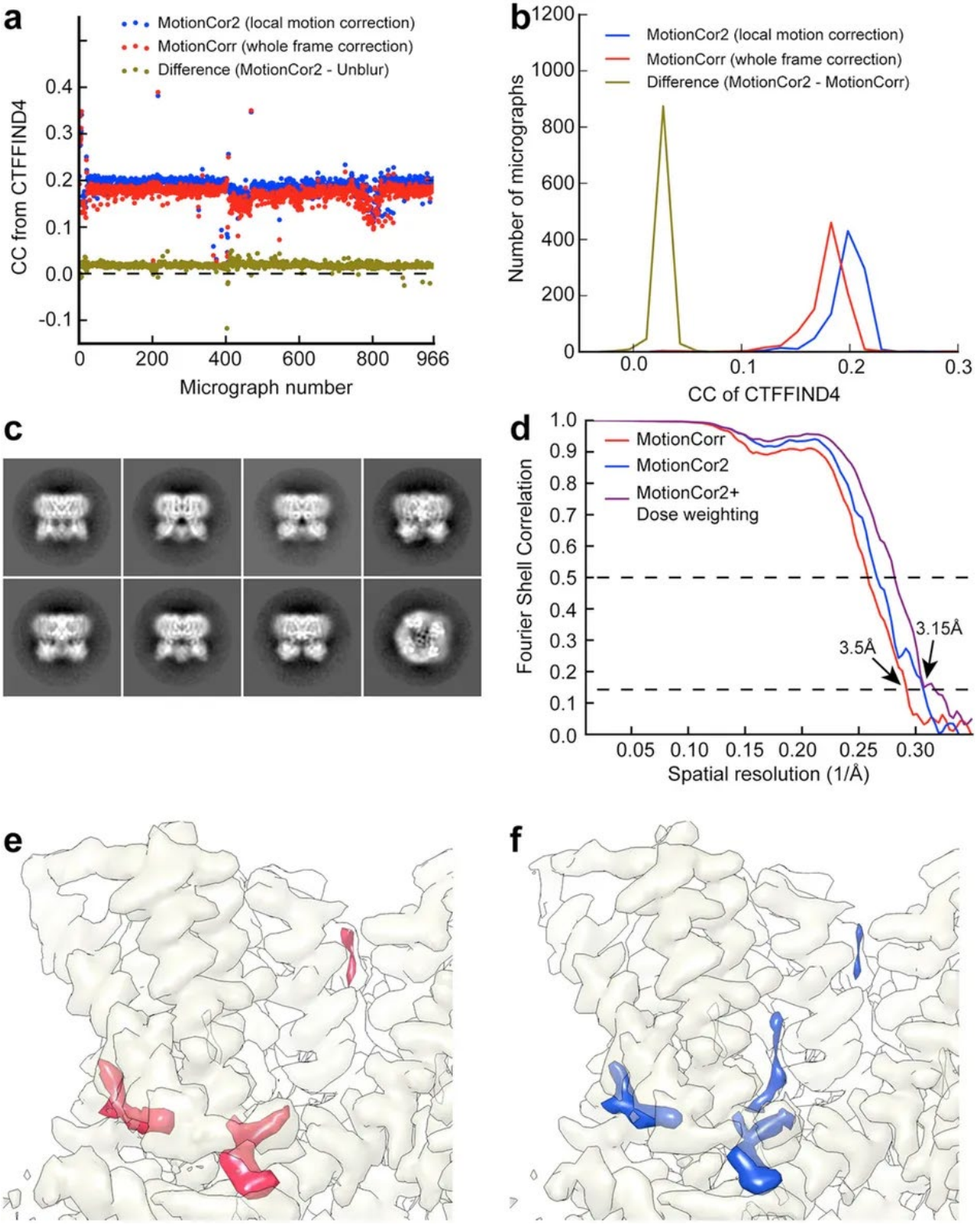
a
MotionCor2 (local motion correction)
MotionCorr (whole frame correction)
Difference (MotionCor2 - Unblur)
CC from CTFFIND4
Micrograph number
b
MotionCor2 (local motion correction)
MotionCorr (whole frame correction)
Difference (MotionCor2 - MotionCorr)
Number of micrographs
CC of CTFFIND4
c
d
MotionCorr
MotionCor2
MotionCor2+
Dose weighting
Fourier Shell Correlation
3.5Å
3.15Å
Spatial resolution (1/Å)
e
f

Figure S1.3: Comparison of motion correction algorithms on TRPV1 (a) Cross Correlation Coefficient of CTFFIND4 from the correction by both MotionCorr (red) and MotionCor2 (blue). (b) Histogram of cross correlation coefficients determined by using CTFFIND4 using image corrected by MotionCor2 (blue) and MotionCorr (red). The difference, which shows the improvement, is shown in brown. (c) Representative 2D class averages of frozen hydrated TRPV1 particles. (d) FSC curves of 3D reconstructions determined from the same dataset after motion correction by MotionCorr (red), MotionCor2 (blue) and MotionCor2 with dose weighting (brown). (e) A representative view of the TRPV1 ion channel generated from previously published density map5. (f) The view of the same region of TRPV1 density map determined after re-process motion correction using MotionCor2. Both maps (e and f) are shown at the same normalized density level, $\sigma = 6$.

We also tested the performance of MotionCor2 with a new dataset of archaeal 20S proteasomes collected from our TF20 microscope operated at 200kV equipped with a Gatan K2 camera. The sub-frame exposure time was set to 75ms, approximately one third of the frame exposure time (200ms) we typically use. The motion captured within each 75ms sub-frame, particularly in the early sub-frames, is significantly less than that captured within a sub-frame with a longer frame exposure time (**Figure S1.4**). About half of these images were recorded with defocus set between 0.4μm to 1.0μm. We picked a small set of 36,000 particles from motion corrected images and determined a 3.4Å reconstruction of the 20S proteasome (**Figure S1.4**). Interestingly, a reconstruction of 4.1Å resolution could be determined using only the 3,297 particles having defocus values less than 1 μm (**Figure S1.4**). In contrast, using the same number of particles, but higher defocus (2-2.5 μm), the resolution of the 3D reconstruction is considerably worse, 4.9Å (**Figure S1.4**). This emphasizes the importance of being able to accurately motion correct very low defocus images.

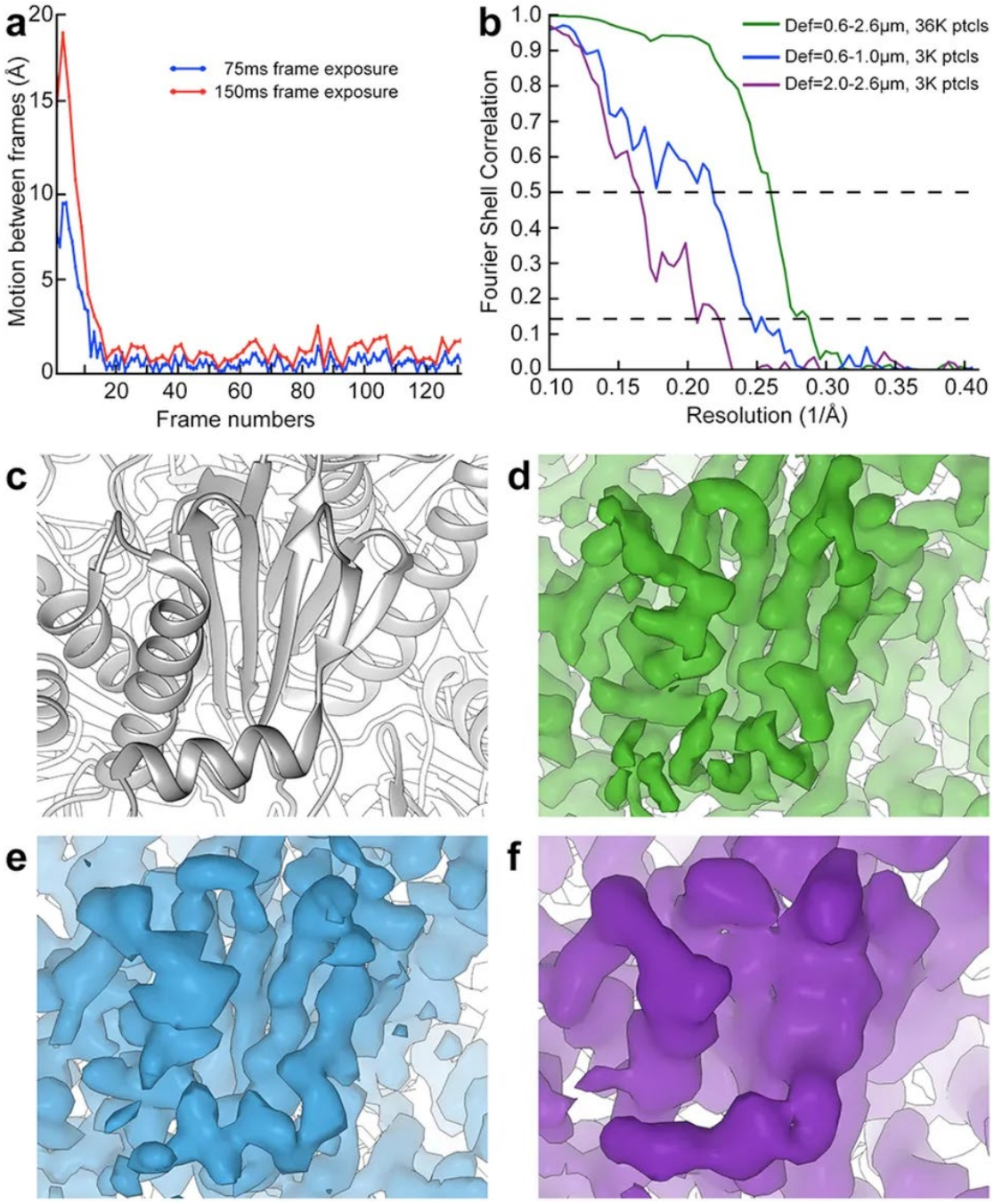

Figure S1.4: Structure of the 20S proteasome on a 200kV electron microscope. (a) A plot of motion between neighboring frames when frame exposure was set to 0.075 second (blue) and 0.15 second (red, by average two adjacent sub-frames to produce a sub-frame that is equivalent to the 0.15 second frame exposure time). (b) Red: FSC curve of an archaeal 20S proteasome 3D reconstruction determined from a dataset of ~36,000 particles collected with a TF20 electron microscope operated at 200kV acceleration voltage. The defocus range was set between 0.6μm and 2.6μm. Blue: FSC curve of a 3D reconstruction using a subset of 3,000 particles with low defocus (0.6μm to 1μm). Purple: FSC curve of a 3D reconstruction of using another subset of 3,000 particles high defocus (2.0μm to 2.6μm). (c) Ribbon diagram of a part of archaeal 20S proteasome. (d) Same region of the 3D reconstruction determined from the entire dataset,

corresponding to red FSC curve in (a). (e) Same region of the 3D reconstruction determined from the subset of 3,000 particles with only low defocused particles (blue FSC curve). (f) Same region of the 3D reconstruction determined from a subset of 3,000 particles with only high defocused particles (purple FSC curve). All maps (d, e and f) are shown at the same density level, $\sigma = 4$.

Discussion

Improving both accuracy and efficiency of motion correction has been an extensive topic of research since direct electron detection cameras and dose-fractionated imaging techniques were first introduced to cryo-EM. Compared with algorithms that only correct the global motion, the new algorithm implemented in MotionCor2 has shown consistent and significant improvements while maintaining superb robustness. Because image motions are corrected at the pixel level prior to any further image processing, it is practical to take advantage of the benefits of dose-weighting without the computationally intensive steps that track and correct motion of individual particles implemented in RELION (Scheres, 2014). Importantly, MotionCor2 is also applicable to tilted data collected by cryo-tomography. Furthermore, by outputting both unweighted and dose-weighted frame averages, CTF determination is optimized. Particles boxed from the weighted images have enhanced low resolution contrast and high-resolution SNR, facilitating more accurate particle alignment and classification, thereby producing improved 3D reconstructions.

Our tests have shown that the new algorithm also works well with dose-fractionated image stacks recorded with defocus and frame exposure times that are significantly lower and shorter than commonly used (**Figure S1.4**). This could potentially impact the collection of single particle cryo-EM datasets when aiming for the highest resolution possible. For atomic-resolution single particle cryo-EM reconstruction, there are many benefits to recording images at low defocus, preferably less than 1um (Wu et al., 2015). Provided that image motion can be corrected accurately, dose weighting allows images to be recorded with sufficiently high total dose to enable

processing while realizing the benefits of low defocus to obtain high resolution. Another practical issue is that motion is most rapid and most anisotropic in the first few subframes, resulting in substantial deterioration of high-resolution information within the first ~3 e-/Å^2 of dose. While these have typically been excluded or substantially down weighted, we have shown that recording images with shorter sub-frame exposure times can reduce such deterioration, again, provided that the local image motion captured in the short sub-frames can be corrected (**Figure S1.4**). Our tests have shown that MotionCor2 provides stable correction of images recorded at both low defocus values and short frame exposure times. Such robust motion correction enabled determination of a 4Å resolution reconstruction using only ~3,000 particles from low-defocused images with short frame exposure times recorded on a TF20 microscope. This dramatically enhances the utility of such an instrument for routine single particle cryo-EM.

Our tests (**Figure 1.1**) suggest that after motion correction by MotionCor2, it is possible to skip the particle tracking part of the RELION polishing procedure, which is by far the most computationally intensive and time consuming step in the refinement of a 3D reconstruction. However, these same experiments suggest that there can be some further benefit from the individual frame weighting component of the RELION polishing procedure. It is perhaps not too surprising that the dose weighting scheme, which was originally implemented in Unblur, may not as accurate as RELION per-frame B-factor weighting, which is determined individually for each dataset. This may also suggest that factors other than pure radiation damage (the model in Unblur) may be affecting the mid and high dose data.

The algorithm implemented in MotionCor2 shares a conceptual link with the unbending procedure used for processing images of two-dimensional crystals (Henderson et al., 1986), which involves local interpolations with a long range constraint imposed by the crystallinity. Similarly,

the x,y, and time-dependent polynomial function used in MotionCor2, also constrains the interpolation, minimizing movements of individual pixels to reduce noise. For typical images of both T20S proteasome and TRPV1, the largest difference of measured shift within a 4,320 x 4,320 Å specimen area is less than ~12Å. Under this estimation, a particle of ~300Å, which is about the size of a eukaryotic ribosome or spliceosome particle, could be distorted by up to ~0.8Å. For very large objects, tracking individual particles might be a better strategy. However, for smaller objects, the accuracy, robustness and computational convenience of MotionCor2 provides a significant practical advantage.

Image processing

Images were subject to motion correction using MotionCorr, Unblur and MotionCor2. After motion correction, images were 2x binned by Fourier cropping implemented in both programs. Un-weighted sums were used for CTF determination using CTFFIND4. Weighted sums were used for automated particle picking using RELION and subsequent image processing. The first sub-frame was excluded in both cases. In MotionCor2, we used per-frame dose of 1.36 e-/Å^2 (20S proteasome), 1.30 e-/Å^2 (TRPV1), and 0.58 e-/Å^2 for the new proteasome dataset collected from TF20 microscope for calculating the critical exposure curves for doseweighting.

Normalized elliptically averaged Fourier spectrum of un-weighted sum of motion corrected sub-frames was used to CTF determination using CTFFIND4. In each dataset, we used 2D class averages calculated from a small number of manually picked particles as references for automatic particle picking using RELION 20. 3D reconstructions of each dataset were also calculated and refined following gold-standard refinement procedure implemented in RELION.

For the 20S proteasome dataset collected at 300kV, a total of 221,623 particles were autopicked from 570 micrographs corrected with MotionCor2. After reference-free 2D classification, particles from classes that did not show high-resolution features were excluded, leaving a final set of 187,011 particles. The coordinates of this particle set were used to extract the equivalent particle images from the Unblur corrected micrographs. Initial reference model for 3D reconstruction and refinement was calculated from the atomic structure of archaeal 20S proteasome low pass filtered to a resolution of 40Å using e2pdb2mrc.py within the EMAN2 package 21. The resolutions of final 3D reconstructions with D7 symmetry are 2.6(9) Å calculated from the dataset corrected by Unblur with dose-weighting, 2.5(7)Å from Unblur followed by particle polishing, 2.5(0)Å from MotionCor2 with dose-weighting, 2.4(6)Å from MotionCor2 with particle polishing and 2.4(6)Å with MotionCor2 with per-frame B-factor weighting respectively. Refined maps were low-pass filtered to the nominal resolution reported and sharpened using the automatic B-factor estimation within the RELION post-processing procedure. Image processing parameters were kept identical between the MotionCor2 and Unblur particle sets such that differences in resulting maps could be attributed primarily to motion correction.

For the TRPV1 dataset, a total of 236,507 particles were auto-picked from 966 micrographs. All particles were used for reference-based 3D classification implemented in RELION, with C4 symmetry applied. Previously determined TRPV1 density map 5 was used as the initial model. A total of 71,825 particles from 3D classes with correct structural features were combined for subsequent 3D refinement. During 3D refinement, we found that applying a soft mask to exclude the density for the cytosolic ankyrin domains during the last five refinement iterations slightly improved the final resolution. Resolutions of final 3D reconstruction calculated

from the MotionCor2 dose-weighted, and unweighted, and original MotionCorr particles are 3.15Å, 3.28Å, and 3.43Å, respectively.

For the 20S proteasome dataset collected with 200kV, image motion was corrected using MotionCor2 only. A total of 36,507 particles were auto-picked and screened by 2D classification from 140 micrographs. Final 3D reconstruction with a D7 symmetry has a resolution of 3.48Å. Furthermore, a 3D reconstruction from a subset of 3,297 particles with defocus between 0.5 – 1.0μm has the resolution of 4.08Å, resolving β-strands within β-sheets. Another 3D reconstruction from a different subset of 3,297 particles with defocus between 2.0 – 2.60μm has the resolution of 4.94Å, where β-strands were not resolved.

References

Abrishami, Vahid, Javier Vargas, Xueming Li, Yifan Cheng, Roberto Marabini, Carlos Óscar Sánchez Sorzano, and José María Carazo. 2015. "Alignment of Direct Detection Device Micrographs Using a Robust Optical Flow Approach." *Journal of Structural Biology* 189 (3): 163–76.

Bai, Xiao-Chen, Israel S. Fernandez, Greg McMullan, and Sjors H. W. Scheres. 2013. "Ribosome Structures to near-Atomic Resolution from Thirty Thousand Cryo-EM Particles." *eLife*. https://doi.org/10.7554/elife.00461.

Bartesaghi, A., D. Matthies, S. Banerjee, A. Merk, and S. Subramaniam. 2014. "Structure of Beta-Galactosidase at 3.2-A Resolution Obtained by Cryo-Electron Microscopy." https://doi.org/10.2210/pdb3j7h/pdb.

Brilot, Axel F., James Z. Chen, Anchi Cheng, Junhua Pan, Stephen C. Harrison, Clinton S. Potter, Bridget Carragher, Richard Henderson, and Nikolaus Grigorieff. 2012. "Beam-Induced Motion of Vitrified Specimen on Holey Carbon Film." *Journal of Structural Biology* 177 (3): 630–37.

Campbell, Melody G., Anchi Cheng, Axel F. Brilot, Arne Moeller, Dmitry Lyumkis, David Veesler, Junhua Pan, et al. (2012). "Movies of Ice-Embedded Particles Enhance Resolution in Electron Cryo-Microscopy." *Structure* 20 (11): 1823–28.

Cheng, Yifan. (2015). "Single-Particle Cryo-EM at Crystallographic Resolution." *Cell* 161 (3): 450–57.

Danev, Radostin, Dimitry Tegunov, and Wolfgang Baumeister. 2017. “Using the Volta Phase Plate with Defocus for Cryo-EM Single Particle Analysis.” *eLife* 6 (January). https://doi.org/10.7554/eLife.23006.

Grant, Timothy, and Nikolaus Grigorieff. 2015. “Measuring the Optimal Exposure for Single Particle Cryo-EM Using a 2.6 Å Reconstruction of Rotavirus VP6.” *eLife* 4 (May): e06980.

Henderson, R., J. M. Baldwin, K. H. Downing, J. Lepault, and F. Zemlin. 1986. “Structure of Purple Membrane from Halobacterium Halobium: Recording, Measurement and Evaluation of Electron Micrographs at 3.5 Å Resolution.” *Ultramicroscopy*. https://doi.org/10.1016/0304-3991(86)90203-2.

Henderson, Richard, and Robert M. Glaeser. 1985. “Quantitative Analysis of Image Contrast in Electron Micrographs of Beam-Sensitive Crystals.” *Ultramicroscopy*. https://doi.org/10.1016/0304-3991(85)90069-5.

Kuhlbrandt, W. 2014. “The Resolution Revolution.” *Science*. https://doi.org/10.1126/science.1251652.

Liao, Maofu, Erhu Cao, David Julius, and Yifan Cheng. 2013. “Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy.” *Nature* 504 (7478): 107–12.

Li, Xueming, Paul Mooney, Shawn Zheng, Christopher R. Booth, Michael B. Braunfeld, Sander Gubbens, David A. Agard, and Yifan Cheng. 2013. “Electron Counting and Beam-Induced Motion Correction Enable near-Atomic-Resolution Single-Particle Cryo-EM.” *Nature Methods* 10 (6): 584–90.

Paulsen, Candice E., Jean-Paul Armache, Yuan Gao, Yifan Cheng, and David Julius. 2015. "Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms." *Nature* 525 (7570): 552.

Rohou, Alexis, and Nikolaus Grigorieff. 2015. "CTFFIND4: Fast and Accurate Defocus Estimation from Electron Micrographs." *Journal of Structural Biology* 192 (2): 216–21.

Rubinstein, John L., and Marcus A. Brubaker. 2015. "Alignment of Cryo-EM Movies of Individual Particles by Optimization of Image Translations." *Journal of Structural Biology* 192 (2): 188–95.

Scheres, Sjors H. W. 2012. "RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination." *Journal of Structural Biology* 180 (3): 519–30.

Scheres, Sjors Hw. 2014. "Beam-Induced Motion Correction for Sub-Megadalton Cryo-EM Particles." *eLife* 3 (August): e03665.

Typke, Dieter, Christopher J. Gilpin, Kenneth H. Downing, and Robert M. Glaeser. 2007. "Stroboscopic Image Capture: Reducing the Dose per Frame by a Factor of 30 Does Not Prevent Beam-Induced Specimen Movement in Paraffin." *Ultramicroscopy* 107 (2-3): 106–15.

Wu, Shenping, Jean-Paul Armache, and Yifan Cheng. 2016. "Single-Particle Cryo-EM Data Acquisition by Using Direct Electron Detection Camera." *Microscopy* 65 (1): 35–41.

Chapter 2 - A simple and robust procedure for preparing graphene-oxide cryo-EM grids

Contributing authors

Eugene Palovcak[1], Feng Wang[1], Shawn Q. Zheng[2], Zanlin Yu[1], Sam Li[1], Miguel Betegon[1], David Bulkley[1], David A. Agard[1,2] and Yifan Cheng[1,2]

[1]Department of Biochemistry and Biophysics, University of California San Francisco, CA 94143
[2]Howard Hughes Medical Institute, University of California San Francisco, San Francisco, CA 94143

Abstract

Graphene oxide (GO) sheets have been used successfully as a supporting substrate film in several recent cryogenic electron-microscopy (cryo-EM) studies of challenging biological macromolecules. However, difficulties in preparing GO-covered holey carbon EM grids have limited their widespread use. Here, we report a simple and robust method for covering holey carbon EM grids with GO sheets and demonstrate that these grids can be used for high-resolution single particle cryo-EM. GO substrates adhere macromolecules, allowing cryo-EM grid preparation with lower specimen concentrations and provide partial protection from the air-water interface. Additionally, the signal of the GO lattice beneath the frozen-hydrated specimen can be discerned in many motion-corrected micrographs, providing a high-resolution fiducial for evaluating beam-induced motion correction.

Introduction

Recent technological breakthroughs have made single particle cryogenic electron microscopy (cryo-EM) a versatile and routine method for structure determination of macromolecules at high-resolution (Bai et al., 2015; Cheng, 2015). With automated data acquisition (Mastronarde, 2005; Suloway et al., 2005) enabled by stable high-end electron microscopes equipped with direct electron detection cameras and streamlined image processing software (Kimanius et al., 2016; Punjani et al., 2017), structure determination by single particle cryo-EM has never been easier. What remains more or less unchanged is the plunge freezing technique (Dubochet et al., 1988), which works well for many samples, particularly structurally stable ones. For fragile complexes, however, preparing good frozen-hydrated cryo-EM grids with intact and monodispersed particles is often a challenging task and a main bottleneck to cryo-EM structure analysis. It has been suggested that exposing protein samples to an air-water interface during plunge freezing can damage fragile protein complexes or induce preferred orientations in thin vitreous ice (Dubochet et al., 1988; Glaeser and Han, 2017; Glaeser, 2018). Additionally, some samples may prefer to stick to the carbon matrix of holey grids instead of being suspended in the vitreous ice spanning the holes (Zhao et al., 2015; Snijder et al., 2017). A common approach to mitigate these problems is to apply a continuous thin layer of substrate to the holey carbon grid, making the distribution of particles in the holes more even and holding protein samples away from the air-water interface (Williams and Glaeser, 1972; Russo and Passmore, 2014; Han et al., 2016; Glaeser, 2018).

Amorphous carbon is the most commonly used substrate (Grassucci et al., 2007), but it adds significant background noise to particle images, limiting its use to relatively large particles such as ribosomes (Gao et al., 2007). Other substrates include monolayer sheets of graphene

(Pantelic et al., 2012; Russo and Passmore, 2014) and two-dimensional crystals of streptavidin (Wang et al., 2008; Han et al., 2017). More recently, thin sheets of graphene oxide (GO) were introduced as a substrate (Pantelic et al., 2010). In two subsequent studies, GO coated grids were used on challenging macromolecular targets to ameliorate a tendency to aggregate at high concentrations and a bad preferred orientation problem, respectively (Bokori-Brown et al., 2016; Boland et al., 2017).

Compared to other options, GO sheets are nearly electron transparent, hydrophilic enough to adhere macromolecules from dilute solutions, inexpensive to purchase or synthesize, and amenable to functionalization (Pantelic et al., 2010; Chen et al., 2012). However, obtaining EM grids evenly covered with one or several layers of GO sheets has not been easy. In our experience, our attempts to use the previously reported drop-casting method (Pantelic et al., 2010) mostly produced grids with irregular coverage of GO sheets over the holes. Typically, fewer than twenty percent of holes were covered by one or a few layers of GO sheets, while the majority of the grids were either covered with multi-sheet aggregates or lacked GO entirely. Because we regularly need to screen tens of cryo-EM grids before finding conditions suitable for automated data acquisition and high-resolution structure determination, reproducibility and ease of manufacture are key considerations for any substrate.

To improve the usable area on GO covered grids, we established a simple and robust surface assembly procedure for evenly covering holey carbon EM grids with one to very few layers of GO sheets. These GO-covered grids are suitable for high-resolution single particle cryo-EM studies of biological macromolecules. We prepared frozen-hydrated archaeal 20S proteasomes using such GO-covered Quantifoil EM grids and collect a dataset resulting in a 2.5Å resolution 3D reconstruction, comparable to our best previous results in freestanding vitreous ice (Li et al.,

2013; Zheng et al., 2017). We also used cryogenic electron tomography experiments (cryo-ET) to confirm that particles are concentrated close to the GO sheet, with about 90% of particles closer to the GO substrate than the air-water interface. In addition, lattice images of the graphene oxide film recorded together with frozen-hydrated 20S proteasome particles provide information for evaluating beam-induced motion correction.

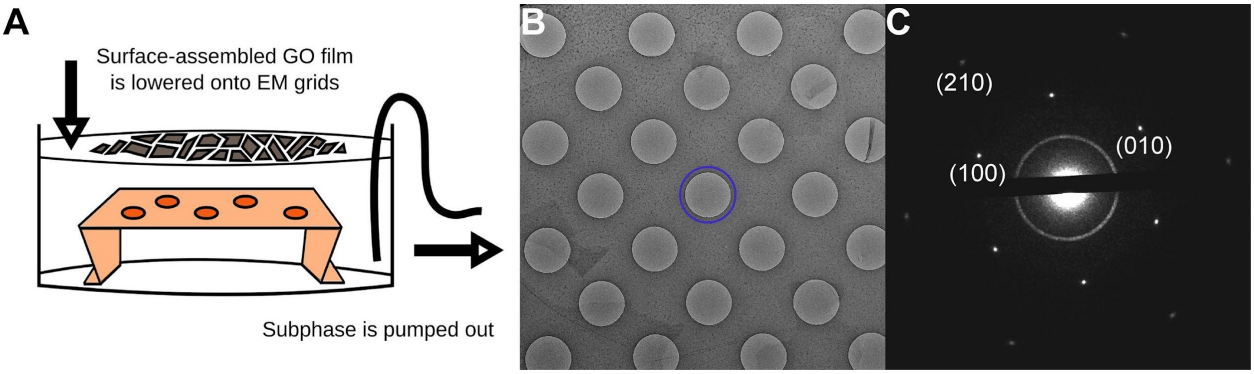

Figure 2.1: EM grids can be evenly coated with GO sheets using surface assembly. (a) Schematic of the apparatus used for surface assembly of thin GO films and subsequent deposition onto holey carbon EM grids. (b) Low magnification image of a holey carbon (Quantifoil) grid covered with a thin film of GO sheets. A hole covered by a single GO sheet is circled in blue. (c) Electron diffraction pattern of the hole in B. Without a scrolled edge, it is hard to discern by image contrast alone whether a hole has a GO sheet spanning it. Instead, the diffraction pattern shows unambiguously the hexagonal lattice of each GO sheet.

Fabricating GO-covered holey carbon cryo-EM grids

Ideally, a GO covered grid would be completely covered in a single monolayer GO sheet without any wrinkled regions or GO aggregates. Commercial GO are available in the form of small flakes. Direct application of an aqueous suspension of GO flakes to a glow-discharged EM grid (drop-casting) (Pantelic et al., 2010) tends to leave many regions of the grid uncovered and deposits high-contrast multi-flake aggregates over many others. Initially we speculated that our commercial GO suspension might have deteriorated with age, so we used bath sonication to break up weakly aggregated sheets followed by centrifugation to isolate mostly large single GO flakes. This treatment reduces the presence of aggregates but does not greatly improve coverage.

GO surface is sufficiently hydrophobic to be enriched at air-water interfaces (Kim et al., 2010). We used this property to assemble a thin, continuous film of GO flakes at the surface of a dish of pure water. By draining the water, the assembled GO film is then slowly lowered onto submerged holey carbon EM grids with their holey carbon film sides facing up (**Figure 2.1**). In agreement with previous reports (Kim et al., 2010; Cote et al., 2011), we found that methanol-dispersed GO flakes spread and float easily on a pure water subphase. Once completely dried, GO flakes stably adhere to holey carbon films and remain adhered during blotting. The boundaries of individual GO flake on EM grids can be directly observed in the transmission electron microscope (TEM) (**Figure 2.1**).

According to the vendors, such as Sigma-Aldrich, their GO flakes are predominantly monolayer sheet, based on atomic force microscope height measurements. Selected area electron diffraction patterns taken from a single flake shows a single hexagonal lattice of Bragg peaks (**Figure 1.1**), suggesting that such flake either contains a single layer of GO sheet, or multiple sheets stacking together coherently as a thin three-dimensional (3D) crystal. Because GO is exfoliated through chemical oxidation of graphite, it possible that some flakes contain more than one layer of GO sheet coherently stack together. We did not attempt to experimentally determine how many layers of GO sheets are coherently stacked together in each flake (Meyer et al., 2007), since the flakes are sufficiently thin that we did not observe any noticeable influence of image contrast. As a result, we assume that most of flakes contain a single layer of GO sheet. Diffraction patterns with sharp high-order spots indicates a long-range periodicity of the GO lattice over the hole (**Figure 2.1**). While GO grids can acquire surface contaminants which render them less hydrophilic, they can be cleaned without damage by brief glow-discharge in air (five to ten

seconds) immediately before use. A detailed protocol for fabricating GO grids by surface assembly is provided in the supplementary methods.

Suspensions of GO sheets can vary significantly in the lateral sheet size and degree of oxidation. The cryo-EM experiments in this study were performed on grids coated with GO sheets from a commercial source (Sigma-Aldrich). We have also recently produced grids with home-made suspensions of GO with larger sheet size. These GO suspensions were made according to an established protocol (Marcano et al., 2010) . Larger GO sheets coat grids more evenly and produce fewer high-contrast 'edges' in images, increasing the efficiency of data collection. Even so, surface assembly is robust to these variations and works well with both commercially-available GO suspensions and home-made GO sheets. The only important parameter to optimize with our protocol is the amount of GO applied to the surface. We typically make one or two test grids and ensure satisfactory coverage by screening in a transmission electron microscope before producing a large batch of grids.

Single particle cryo-EM of archaeal 20S proteasome on a GO grid

Using the archaeal 20S proteasome as a test specimen, we evaluated the practicality of using GO grids for high-resolution single particle cryo-EM. As previously reported, the commonly used plunge freezing procedure works well for GO grids, though we have found that longer blotting times (10-30 seconds) are often preferred. For the 20S proteasome, achieving an optimal particle distribution required a specimen concentration approximately ten times lower with GO (0.05 mg/mL) than without (0.5 mg/mL) (**Figure 2.2, S2.1**). By cryo-ET, we confirmed that 20S proteasome particles directly adhere to the GO substrate, as almost 90% of particles were closer to the GO substrate than the air-water interface (**Figure 2.2**). Based on the locations of the few

proteasomes and gold nanoparticle fiducials not adhered to the GO face, we estimate the thickness of the vitreous ice in this hole to be about 65 nm, while the deviation in the Z-location of 3D template-matched 20S proteasome particles is only about 5 nm. This suggests that 20S proteasomes on GO are protected from the air-water interface and lie on a common plane, a clear benefit when estimating micrograph defocus.

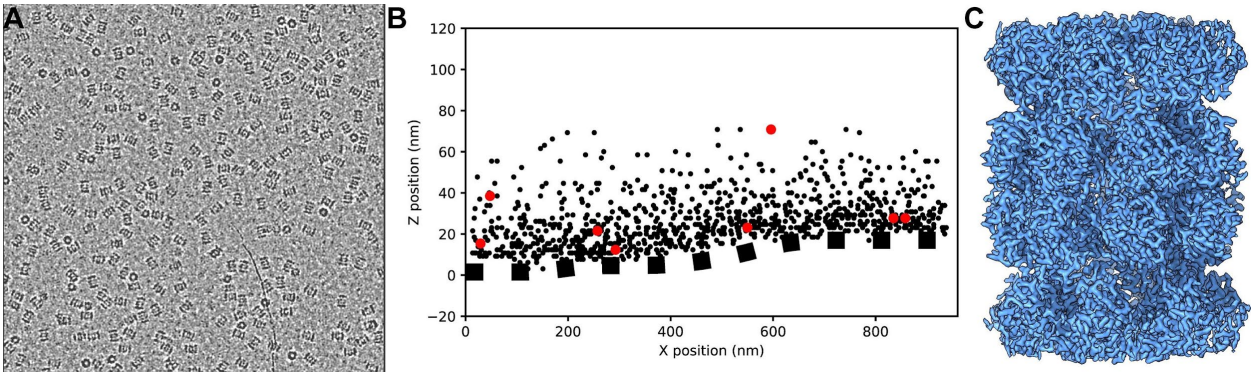

Figure 2.2: High-resolution single-particle cryo-EM on GO grids. (a) Micrograph of 20S proteasome particles over a GO support film. The GO film adds minimal background contrast. The specimen was applied at a concentration 0.05 mg/mL, approximately ten times less than we typically use for this specimen. (b) 3D localization of 20S particles from cryogenic electron tomography (cryo-ET). This tomogram was taken from a different GO grid than the one we used for the single-particle results: the 20S proteasome concentration here was higher and the GO substrate was far thicker (5–8 sheets). These conditions enabled us to identify how densely the GO surface could be coated with macromolecular specimen and to discern the position of the GO layer within low-resolution tomogram. Fiducial markers and points across the GO substrate surface were picked manually, while 20S proteasome particles were located in the tomogram using template matching. The GO is not flat but ripples and bends, as shown by the dotted line fitted by a spline to the points manually picked on the GO surface. The maximal distance between a picked proteasome particle and the GO substrate surface was 66 nm, bounding the relative position of the air-water interface. 862 out of 969 picked proteasome particles (88.9%) are closer to the GO substrate than the air-water interface, strongly suggesting a direct interaction. All tomographic processing was performed using the IMOD software suite (Mastronarde, 2005). (c) 3D reconstruction of the 20S proteasome on a GO support resolved at 2.5Å.

To compute a 3D reconstruction by single-particle analysis, we used the SerialEM automated data acquisition procedure to collect 740 dose-fractionated stacks. When choosing areas for data acquisition, we did not attempt to distinguish if holes were covered by GO or not. In this

dataset, approximately 73% of the collected images were of high-quality, judged from the particle distribution and lack of numerous high contrast GO edges. 10% of micrographs had thick crystalline ice quality likely unrelated to the GO, 10% of the holes are missing the GO support, and 7% of the holes had suboptimal coverage (usually too many overlapping sheets). Beam-induced motion was corrected using MotionCor2 (Zheng et al., 2017).

For 3D reconstruction, we included only particles collected from holes covered with GO sheets and did not attempt to computationally remove the periodic graphene lattice, as might be done when using other crystalline lattice supports such as streptavidin crystal grids (Wang et al., 2008; Han et al., 2016). From this dataset, we determined a 3D reconstruction of the 20S proteasome at a resolution of 2.5Å (**Figure 2.2**). The density map is visually indistinguishable in quality from our previously reported map of the 20S proteasome in freestanding vitreous ice but used fewer particles in the final reconstruction (117,578 particles on GO vs. 187,011 in freestanding vitreous ice (Zheng et al., 2017). This suggests that despite the added background contrast and the micrographs lost due to defects in the GO grid fabrication, GO poses no significant barrier to achieving high-resolution reconstructions.

Evaluating correction of beam-induced motion

GO sheet typically maintains a long-range periodicity (**Figure 2.1**) that can tolerate the amount of electron beam radiation used to image the biological sample (**Figure S2.2**). For a typical image recorded as a movie stack of subframes, only weak spots are seen in the sum image before motion correction (**Figure 2.3**). These peaks can also be observed in short three-frame averages during the exposure, suggesting that the GO lattice is not significantly deteriorated by radiation damage but is primarily blurred by beam-induced motion. After full motion correction with

MotionCor2 (Zheng et al., 2017), the power spectrum show one complete set of hexagonally arranged peaks around 2.1Å resolution, beyond the physical Nyquist frequency of the K2 camera (2.43Å at the selected magnification). These peaks signal the first order reflections of the underlying GO lattice, with each hexagonal set of peaks corresponding to a single GO sheet. Reflections are typically very sharp, with an average half width of 10 Fourier pixels, suggesting that the vitrified GO lattice covering the image area maintains a long-range periodicity.

In MotionCor2 (Zheng et al., 2017), global motions are corrected by iteratively refining translational shifts for each frame, while nonuniform local motions are corrected by fitting numerous trajectories of local motions to a time-variant polynomial function. This polynomial estimates the instantaneous shift for any location in the image at any time point in the exposure, allowing local motions to be corrected smoothly at every image pixel. The number of image patches where local motion trajectories are measured is a free parameter in MotionCor2: a coarse grid of patches (i.e. 3x3 patches) will provide fewer measurements, but they should have high signal-to-noise ratio (SNR), while a finer grid (9x9 patches) will provide a finer measurement of localized motion, but at the expense of increased computation and a lower per-patch SNR.

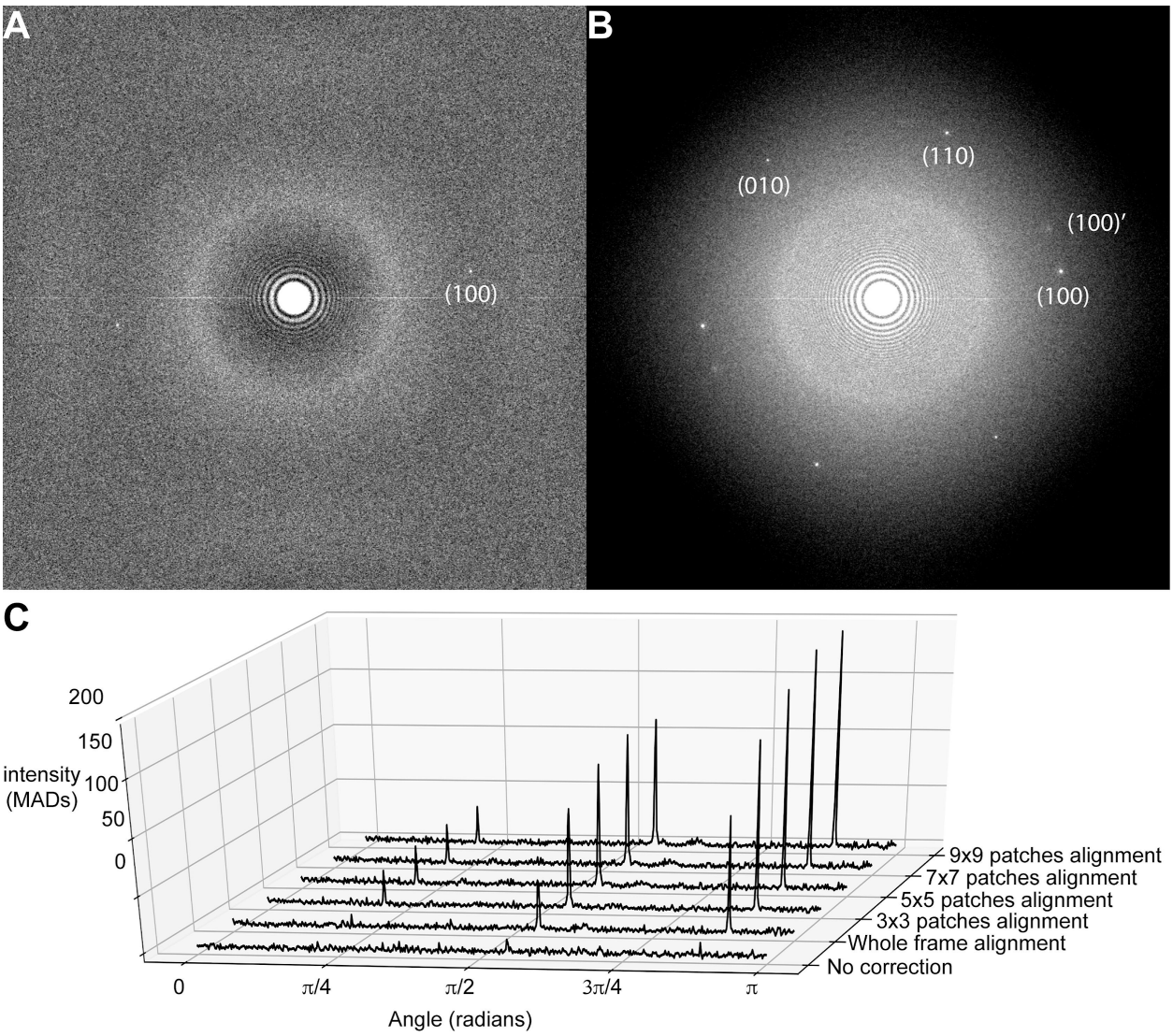

Figure 2.3: Evaluation of motion correction using GO peak intensities. Power spectra from the image in Fig. 2A before (a) and after (b) motion correction. Power spectra were calculated by averaging the periodograms of overlapping 512 × 512 pixel windows using a script in Python. (c) GO peak height increases as the number of local motion patch trajectories calculated by MotionCor2 is increased. To generate 1D radial profiles showing the GO peaks with respect to the background, the aforementioned power spectrum was transformed into polar coordinates by cubic spline interpolation and a radial band of intensities from 2.0 Å to 2.2 Å was extracted. The maximum intensity component in this band was taken for each angle sampled in the half-circle 0–π. This was done because the actual spatial frequency of the GO peak deviated from the expected position at 2.13 Å, likely due to residual specimen tilt or uncorrected anisotropy in the magnification system (Zhao et al., 2015). To compare 1D profiles between different motion correction schemes, the intensities in each 1D profile were scaled according to the median absolute deviation (MAD), which is a robust measure of scale. All calculations were performed in python with the scientific computing library scipy, with code available on request.

GO peaks are nearly absent in the power spectra before motion correction, but they are readily visualized after global motion correction (**Figure 2.3**). Applying local motion correction, we observe a trend where increasing the number of patches significantly increases GO peak intensity. For this particular image, using 7x7 patches increased the peak intensities approximately 200% with respect to the global alignment, while 9x9 patches provided no additional significant increase. While the intensities of all three distinct GO lattice peaks increase with better motion correction, the amount of improvement is different for each. We speculate that this could occur if the image still contains uncorrected motion, such as that occurring within each frame and if this motion were orthogonal to the direction of the foreshortened peaks.

Recently, Han et al (2017) observed a correlation between the resolutions of streptavidin 2D crystals and of the reconstruction determined from particles absorbed to the streptavidin substrate (Han et al., 2017) and proposed the use of peak intensity from a 2D streptavidin crystal to evaluate cryo-EM image quality. Resolution of high-resolution peaks from streptavidin crystal are sensitive indicators of specimen damage incurred during vitrification or exposure to the electron beam, and can thus be used to exclude images of damaged specimen.

Compared with a protein crystal lattice, GO is tolerant to vitrification and electron radiation damage, maintaining long-range order throughout the exposure (**Figure S2.2**). Beam-induced motion is the primary factor that weakens GO reflections during an exposure. The intensity and sharpness of the reflections after motion correction could be directly correlated with the quality of the correction.

To use resolutions or peak intensities of substrate crystals to evaluate corrections of local motions, the larger unit cell of streptavidin lattice is more advantageous. However, unlike GO lattice whose crystallinity can be verified by electron diffractions, crystallinity of streptavidin 2D

crystals varies and image unbending can also improve the resolution (Henderson et al., 1986), making it necessary to separate resolution improvement by unbending or motion correction.

Discussion

We have established a simple and robust procedure for covering holey carbon EM grids with thin films of GO sheets and have demonstrated the utility of the resultant grids for determining 3D reconstructions of macromolecules at high-resolution. We also showed that GO peaks in the image power spectrum are sensitive fiducial markers for the quality of the beam-induced motion correction. By simplifying their production, we anticipate that our method will make the benefits of GO grids more widely accessible.

Nevertheless, we believe the GO grids described here are only one step towards a more universal substrate for high-resolution single-particle cryo-EM. In the course of testing this protocol, we have attempted to use GO grids in several challenging, on-going cryo-EM projects. In some cases, macromolecular particles were bound and concentrated onto the GO surface in a native state like we report here for the 20S proteasome (**Figure 2.2B**). In other cases, particles were not visibly bound to the GO or appeared denatured (data not shown). This could occur if the proteins have no affinity for GO, as they would then have no protection from the air-water interface. There have also been reports that GO itself can destabilize protein structure and compromise enzymatic activity (Bai et al., 2017). In either case, an obvious solution is to functionalize the GO surface, increasing its affinity for interactions that preserve protein structure while passivating residual non-oxidized hydrophobic domains that may promote denaturation. GO has abundant epoxide groups on its surface that are amenable to such functionalization. We are actively working along these lines to improve the applicability of GO grids to fragile specimens.

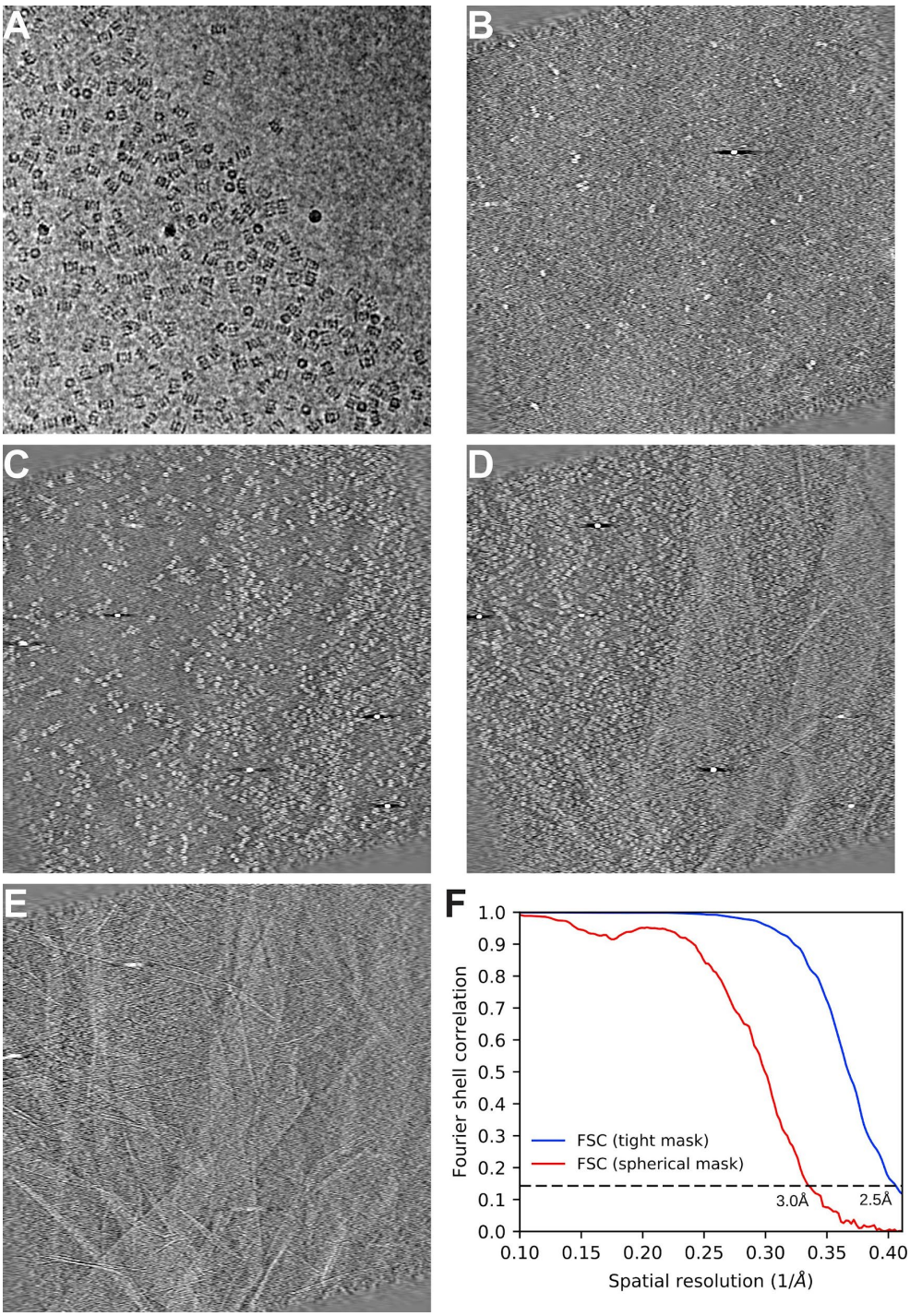

Figure S2.1: Cryo-EM of proteasome particles prepared using GO covered grid. (a). Image of a hole partially covered with a GO. Fewer proteasomes are observed in the upper right-hand corner where the GO support is absent. The additional background contrast added by the GO substrate is barely perceptible. B-D. Projections through Z-slices of the tomogram summarized in Figure 2B. The slices depict (b). the air-water interface (58-80 nm away from the lowest apparent point of the GO substrate), (c) the higher part of the GO surface where bound proteasomes are observed (23-39nm), and (d), the lower part of the GO surface

where bound proteasomes are observed (8-23nm). (e). Gold-standard FSC curves for the 20S proteasome, calculated from independently reconstructed half-maps.

If a protein specimen is properly adhered to GO grids, may be better than if no substrate is used, as particles common surface should have locally-correlated heights, improving the accuracy of defocus estimation. GO peaks in the power spectrum do not follow precise hexagonal symmetry, and that the lattice spacing for each of the three peaks vary slightly around 2.1Å. This can be caused either by anisotropic magnification of the microscope or a slight tilt of the specimen, or a combination of the both. If the anisotropic magnification of the microscope is precisely calibrated, it is possible to derive the tilt angle and orientation of the specimen, facilitating a better local defocus determination. At a resolution of ~2.5Å, we did not notice any influence of the underlying GO lattice on particle alignment or classification for the 20S proteasome reconstruction. We therefore did not attempt to computationally remove the underlying GO lattice. It may become necessary to do so if the resolution of a reconstruction is beyond the GO lattice space, ~2.1 Å.

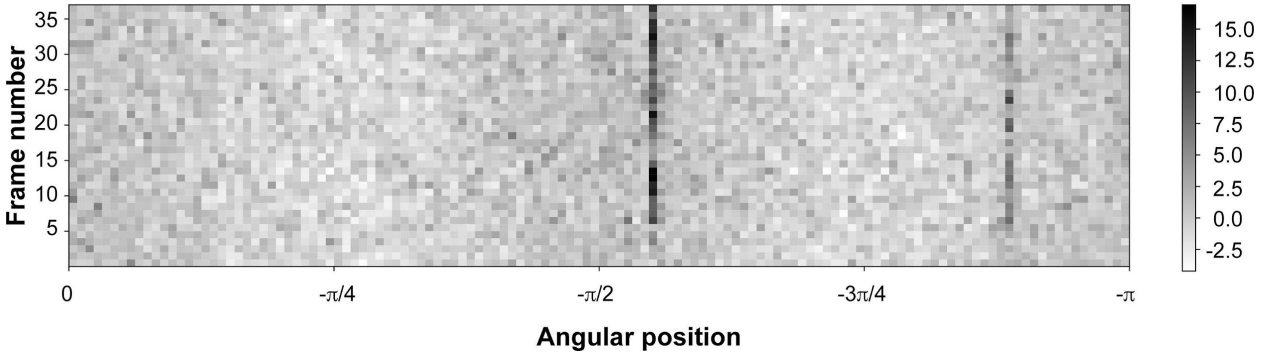

Figure S2.2: From the dose-fractionated stack used in Figure 2.3, we calculated power spectra from running three-subframe averages of without motion correction. Each row shows the 1D radial profile of a power spectrum calculated from an averaged image of three consecutive subframes. The 1D radial profile was calculated identically to those in Figure 2.3C. Two out of three GO peaks are visible in most three-frame average images. The third peak is too weak to be detected from power spectrum of individual frames. While the peaks are not detected at the beginning of the exposure when beam-induced motion is greatest, they are not obviously reduced at the end of the exposure. The color scale represents the MAD.

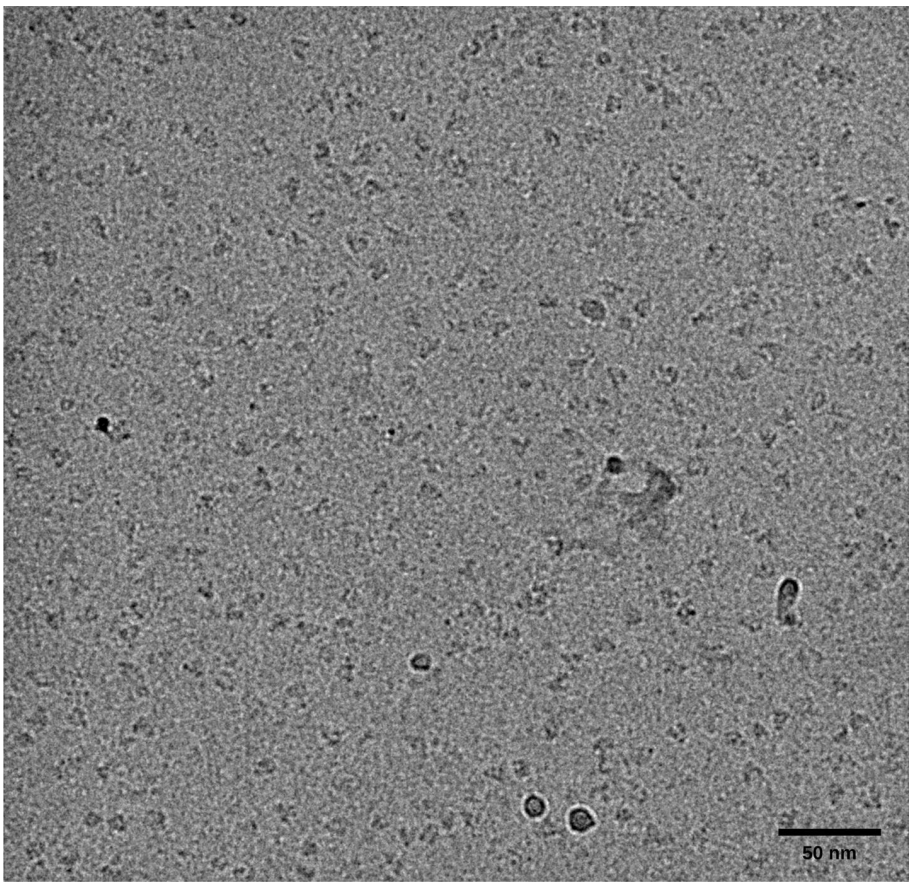

Figure S2.3: Image of a 170kDa complex on GO. Region of a cryo-EM image of a substrate-bound E. coli HptG complex (expected molecular weight ~170kDa) weakly crosslinked with bissulfosuccinimidyl suberate and bound to a GO grid. This particle is about 4 times smaller than the 20S proteasome and was never observed in vitreous ice without a support film, even at very high concentrations. On GO grids, individual particles are apparent when the specimen is applied at 100nM, suggesting that the GO is binding and concentrating the complex onto its surface. For this experiment, we used a GO grid coated with a homemade GO suspension. The specimen was applied to a GO grid, blotted for 4s on a Mark IV Vitrobot (100% humidity, blot force 0), and plunged into liquid ethane. The image was collected on a 200kV Talos Arctica equipped with a Gatan K2 camera (1.15Å/pixel, 60e- total exposure, 0.2s/frame, 12s exposure) and motion corrected with MotionCor2 (Zheng et al; 2017). The defocus of the image is estimated to be 2.4 microns (Zhang, 2016).

Finally, we showed that subtle improvements in beam-induced motion correction can be detected by examining the intensity of GO peaks present in the power spectra of differently motion corrected micrographs. Note that we did not use GO peak intensity as an objective function to directly optimize subframe alignment, in which the motions between consecutive subframes larger than the unit cell of the GO lattice (2.13Å) could degenerate subframe registration solutions (i.e.,

aliasing). However, it is possible to optimize the motion correction by maximizing the GO peak intensity as an objective function. This is because the residual local motions after whole-frame rigid body motion correction are relative small (Supplementary Figure 1B in Zheng et al, 2017), thus a displacement between two consecutive subframes is likely much less than an angstrom and registration of shifts larger than the GO lattice unit cell can be excluded. If the periodic GO signal is used only to locally refine the initial registration trajectory estimated from aperiodic low-resolution image features (as MotionCor2 already does), deleterious artifacts of overfitting will be unlikely, permitting faithful recovery of the highest resolution features captured in the dose-fractionated electron micrographs.

Methods and Materials

Graphene oxide (GO) grids were made on Quantifoil 200 mesh R 1.2/1.3 holey carbon grids by the method described in the Supplementary Protocol. The GO suspension was purchased from Sigma Aldrich (777676 ALDRICH). We have also successfully produced GO grids using home-made suspensions of GO with larger lateral sheet sizes. For this we used the improved method of synthesizing GO from James Tour's lab, which is described in detail in Marcana et al. (2010). GO grids were used for cryo-EM once they had completely dried, several hours after they were made.

Thermoplasma acidophilum 20S proteasomes were expressed and purified as previously described (Li et al., 2013) and stored in aliquots at -80^{o}C. An aliquot of 20S proteasome was diluted to 0.05 mg/mL in buffer containing 25mM Tris pH=7.5 and 150mM NaCl. 2.5uL of 20S proteasome was applied to a GO grid and allowed to incubate for 30S in the chamber of a Mark

III Vitrobot at 100% humidity. The specimen was blotted for 6s with Whatman #1 filter paper using an offset of 0mm and plunged into liquid ethane.

Cryo-EM grids were loaded into a TF30 Polara microscope equipped with a Gatan K2 camera operated in super-resolution mode. Automated data acquisition was performed with SerialEM and 740 micrographs were corrected. Beam-induced motion was corrected in MotionCor2 (Zheng et al., 2017). Per-frame electron dose weighting was performed with MotionCor2 using a nominal electron dose of 1.2 e^-/A^2*frame and the radiation damage model derived in (Grant and Grigorieff, 2015). Micrographs were visually inspected and segregated into optimal GO coated (73%), overcoated GO (10%), poor ice (10%), or uncoated (7%) images. 540 micrographs with optimal GO coating were selected for continued processing. CTF parameters were estimated with Gctf (Zhang, 2016) and particles were picked with Gautomatch (https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) using a template of a 20S proteasome side view. Mispicked particles were removed after 2D classification in cryoSPARC (Punjani et al., 2017)and subjected to homogeneous 3D refinement with D7 symmetry. Masking and FSC estimation were performed automatically in cryoSPARC. The resulting map was compared visually to previous 20S proteasomes maps from our lab with UCSF Chimera.

References

Bai, X.C., McMullan, G., and Scheres, S.H. (2015). How cryo-EM is revolutionizing structural biology. Trends Biochem Sci *40*, 49-57.

Bokori-Brown, M., Martin, T.G., Naylor, C.E., Basak, A.K., Titball, R.W., and Savva, C.G. (2016). Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat Commun *7*, 11293.

Boland, A., Martin, T.G., Zhang, Z., Yang, J., Bai, X.C., Chang, L., Scheres, S.H., and Barford, D. (2017). Cryo-EM structure of a metazoan separase-securin complex at near-atomic resolution. Nat Struct Mol Biol *24*, 414-418.

Chen, D., Feng, H., and Li, J. (2012). Graphene oxide: preparation, functionalization, and electrochemical applications. Chem Rev *112*, 6027-6053.

Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. Cell *161*, 450-457.

Cote, L.J., Kim, J., Tung, V.C., Luo, J.Y., Kim, F., and Huang, J.X. (2011). Graphene oxide as surfactant sheets. Pure Appl Chem *83*, 95-110.

Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. Q Rev Biophys *21*, 129-228.

Glaeser, R.M. (2018). Proteins, interfaces, and cryo-EM grids. Current opinion in colloid and interface science *https://doi.org/10.1016/j.cocis.2017.12.009*.

Glaeser, R.M., and Han, B.G. (2017). Opinion: hazards faced by macromolecules when confined to thin aqueous films. Biophys Rep *3*, 1-7.

Han, B.G., Watson, Z., Cate, J.H., and Glaeser, R.M. (2017). Monolayer-crystal streptavidin support films provide an internal standard of cryo-EM image quality. J Struct Biol.

Han, B.G., Watson, Z., Kang, H., Pulk, A., Downing, K.H., Cate, J., and Glaeser, R.M. (2016). Long shelf-life streptavidin support-films suitable for electron microscopy of biological macromolecules. J Struct Biol *195*, 238-244.

Henderson, R., Baldwin, J.M., Downing, K.H., Lepault, J., and Zemlin, F. (1986). Structure of purple membrane from halobacterium halobium: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. ultramicroscopy *19*, 147-178.

Kim, J., Cote, L.J., Kim, F., Yuan, W., Shull, K.R., and Huang, J.X. (2010). Graphene Oxide Sheets at Interfaces. J Am Chem Soc *132*, 8180-8186.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife *5*.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods *10*, 584-590.

Marcano, D.C., Kosynkin, D.V., Berlin, J.M., Sinitskii, A., Sun, Z., Slesarev, A., Alemany, L.B., Lu, W., and Tour, J.M. (2010). Improved synthesis of graphene oxide. ACS Nano *4*, 4806-4814.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol *152*, 36-51.

Meyer, J.C., Geim, A.K., Katsnelson, M.I., Novoselov, K.S., Booth, T.J., and Roth, S. (2007). The structure of suspended graphene sheets. Nature *446*, 60-63.

Pantelic, R.S., Meyer, J.C., Kaiser, U., Baumeister, W., and Plitzko, J.M. (2010). Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. J Struct Biol *170*, 152-156.

Pantelic, R.S., Meyer, J.C., Kaiser, U., and Stahlberg, H. (2012). The application of graphene as a sample support in transmission electron microscopy. Solid State Communications *152*, 1375-1382.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296.

Russo, C.J., and Passmore, L.A. (2014). Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. Nat Methods *11*, 649-652.

Snijder, J., Borst, A.J., Dosey, A., Walls, A.C., Burrell, A., Reddy, V.S., Kollman, J.M., and Veesler, D. (2017). Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. J Struct Biol *198*, 38-42.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. J Struct Biol *151*, 41-60.

Wang, L., Ounjai, P., and Sigworth, F.J. (2008). Streptavidin crystals as nanostructured supports and image-calibration references for cryo-EM data collection. J Struct Biol *164*, 190-198.

Williams, R.C., and Glaeser, R.M. (1972). Ultrathin carbon support films for electron microscopy. Science *175*, 1000-1001.

Zhao, M., Wu, S., Zhou, Q., Vivona, S., Cipriano, D.J., Cheng, Y., and Brunger, A.T. (2015). Mechanistic insights into the recycling machine of the SNARE complex. Nature *518*, 61-67.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332.

Chapter 3 - Nucleic acid grids for stabilizing chromatin specimens during vitrification

Contributing Authors

Eugene Palovcak[1], Un Seng Chio[1], Nathan Gamarra[1], Elise Munoz[1], Laura Hsieh[1], Anton A. A. Smith[2], Geeta Narlikar[1], Yifan Cheng[1,2]

[1]Department of Biochemistry and Biophysics, University of California San Francisco, CA 94143
[2]Department of Materials Science & Engineering, Stanford University, Stanford, California
[3]Howard Hughes Medical Institute, University of California San Francisco, San Francisco, CA 94143

Preface

This chapter describes my efforts with Un Seng Chio and Geeta Narlikar's lab to develop a general and effective method for stabilizing chromatin specimens for single-particle cryo-EM. It is a direct extension of the graphene-oxide grid development described in chapter 2. This research is not finished, as there are still open questions about the interpretations of the structures we obtained, but I feel that my role in the work has been highly successful and worthy of inclusion in this thesis. This preface is an account of the research and experiments that led up to the preliminary results obtained for this project.

Chromatin is the packaged state of the eukaryotic genome and consists of long, linear double-stranded DNA chromosomes wrapped around highly-charged disc-shaped protein complexes called histones. While a large part of chromatin's role is to organize and coil the 3 billion DNA base-pairs of the genome into a micron-sized nucleus, it also serves an important

regulatory role, selectively and dynamically masking the binding sites of sequence-specific DNA binding proteins. In fact, the function of many transcription factors is to recruit large multisubunit enzyme complexes that chemically modify or physically displace the histone proteins at genes or regulatory sites (termed chromatin modifying-complexes or chromatin remodelers, respectively), enabling the orchestrated activation or silencing of gene expression. For this reason, chromatin modifying-complexes and chromatin remodelers play key roles in the programming of cellular state and mutations in their proteins are very common in complex human diseases such as cancer, mental illness, and developmental genetic illnesses. Understanding the function of these complex enzymes might enable novel, rational pharmacological interventions in these diseases.

Like much of the enzymology, the biochemical study of chromatin modifying enzymes and chromatin remodelers would benefit from high-resolution structural characterization, preferably on native chromatin substrates. The minimal constituent of chromatin is a histone octamer wrapped 2.5x by 160 base-pairs of double-stranded DNA, termed a nucleosome. The structure of a nucleosome core particle (a nucleosome without any 'linker' DNA at the entry or exit sites), was determined in 1997 by Karolin Luger. However, despite decades of work, subsequent efforts to crystallize nucleosome-protein complexes have not been very successful. Additionally, modern single-particle cryo-EM has underperformed on chromatin specimens compared to other classes of previously challenging specimens such as membrane proteins, ribosomes, and filaments.

While mostly unstated in the literature, when I joined Yifan's lab it was common knowledge in the cryo-EM community that chromatin specimens do not survive the blotting and vitrification steps required for preparing thin films of frozen-hydrated specimen. Without such thin electron-transparent films of specimen, it is impossible to take cryo-EM images. Our lab had first-

hand experience in this problem from our long-standing collaboration with Geeta Narlikar, one of the world's preeminent chromatin enzymologists.

Geeta's lab had been studying a human chromatin remodeling enzyme called snf2h. Using elegant biochemical and spectroscopic experiments, the Narlikar lab was able to show that snf2h acts as a dimer to evenly space linear arrays of nucleosomes, promoting a condensed, repressive state of chromatin. Intriguingly, they also showed that this activity requires deformations of the histone octamer, which was previously assumed to act as a rigid, conformationally inert substrate. Geeta and Yifan wanted to see these results validated and further illuminated with high-resolution 3D structures of the relevant complexes, but despite years of effort and hundreds of hours of data collection, only one high-resolution dataset of snf2h bound a nucleosome was ever obtained. It was not clear why the particles in this dataset were not denatured during vitrification, and so the results could not be reproduced.

I became interested in the problem of stabilizing fragile, denaturation-prone macromolecular specimens during vitrification. At Berkeley, Bob Glaeser had been talking for several years about the role the air-water interface in this process. He claimed that when a film of a protein-containing solution is blotted and thinned, the surface-to-volume ratio increases significantly and the probability that a macromolecule will collide with the air-water interface becomes very high, essentially inevitable. Air is very hydrophobic and air-water interfaces are known to unfold proteins via the phenomenon of surface denaturation. Bob's strategy for avoiding the surface denaturation during vitrification was to engineer affinity grids that would bind macromolecules to an electron-transparent substrate and hold them away from the air-water interface. His method involved coating grids with large, 2D crystals of streptavidin and randomly biotinylating the macromolecular specimens. After attempting his method several times, I found it

impractical: the streptavidin crystal layer was difficult to produce and produced a significant background in images. Even so, the logic of the approach was very appealing. This led me to develop the robust procedure for producing graphene-oxide grids described in chapter 2.

Unmodified GO grids turned out to be mostly ineffective substrates for single particle cryo-EM. Despite being heavily oxidized, GO is quite hydrophobic. Maintaining a thin frozen-hydrated film was difficult on GO grids, and most grids would appear 'dry' when examined in the electron microscope, as if all the liquid either pooled up around the grid bars or was blotted away during specimen preparation. Additionally, many specimens had no affinity for unmodified GO. The obvious solution was to modify the surface of the GO grid to bind cryo-EM specimens. In this regard, GO is an especially appealing substrate, as its surface is coated with reactive epoxide groups and its functionalization is well-described in the literature. Feng Wang, a material scientist recruited by David Agard to work on GO grids, discovered that treatment of GO grids with an aprotic solution of diamine produced hydrophilic GO grids (presumably through nucleophilic attack of the epoxide). Feng's GO-amine grids bound many macromolecular specimens away from the air-water interface without imposing a strong preferred orientation. However, GO-amine grids were strongly positively charged, and whenever I attempted to apply chromatin specimens to them, I observed denatured nucleosomes. This is consistent with the positively charged surface of the GO-amine grid competing with the histone octamer proteins for the anionic phosphate backbone of the nucleosomal DNA.

The initial plan developed with Elise Munoz in Geeta's lab was to functionalize the GO grids with single-stranded DNA oligonucleotides using the same nucleophile-epoxide chemistry used for the GO-amine grids. Then, we would produce nucleosomes with a complementary single-stranded overhang on one of the linkers. These complementary single-stranded oligonucleotides

would hybridize, tethering nucleosomes and any bound protein complexes to the surface of the GO grid. We quickly determined that ssDNA-GO grids could indeed bind and stabilize chromatin specimens through vitrification under conditions that would usually denature them. However, we also discovered that the complementary overhang she had engineered onto the nucleosomes was not necessary: any nucleosome seemed to bind the grids, so long as we included divalent cations in the buffer. This was fortunate because installing a single-stranded linker onto a nucleosome with otherwise double-stranded DNA was biochemically laborious. We reasoned that in the presence of divalent cations (e.g. $Mg2^{+}$), ionic crosslinking could occur between the phosphate backbone of the GO-grafted ssDNA and the phosphate backbone of the nucleosomal DNA. This produced a relatively high-affinity yet non-specific interaction between the ssDNA-GO grid and nucleosome-enzyme complexes.

In the rest of this section, I describe the results I obtained with Un Seng Chio, a joint post-doc in the Narlikar and Cheng labs, using ssDNA-GO grids to solve several high-resolution structures of the snf2h-nucleosome complex in the ADP-BeFx state, finally reproducing the structure from the single good dataset collected in the previous years. Because we were able to obtain a large amount of high-quality particle images of uncrosslinked snf2h-nucleosome complexes, we were also able to apply novel 3D variance analyses to resolve conformational changes in the nucleosome that we tentatively believe are consistent with the snfh2-induced octamer deformations previously observed jointly by the Gross and Narlikar labs with methyl-TROSY NMR spectroscopy.

Nucleic acid grids stabilize chromatin remodeler complexes during vitrification

To stabilize complexes of nucleosomes and chromatin remodelers for cryo-EM imaging, we developed nucleic acid affinity grids. The general principle is to coat an EM grid with an electron-transparent substrate functionalized with single-stranded nucleic acids. In the presence of divalent cations, polyanionic polymers are known to form weak ionic crosslinks. We reasoned that such interactions would be ideal for binding chromatin complexes because they would not be strong enough to compete with histones or nucleosome-binding proteins for interactions with the nucleosomal DNA and they would be nonspecific and abundant enough to produce a relatively random distribution of binding orientations on the support film (**Figure 3.1**). GO-grids were functionalized 5'-labeled single-stranded DNA oligonucleotides (NH_2-ssDNA) by incubating a GO grid in an aprotic solution of the oligos overnight at room temperature (see Materials and Methods). This resulted in highly hydrophilic grids after extensive rinsing and drying, consistent with covalent functionalization of the GO surface by polyanionic ssDNA, presumably via attack of GO surface epoxides by the nucleophilic 5' amine.

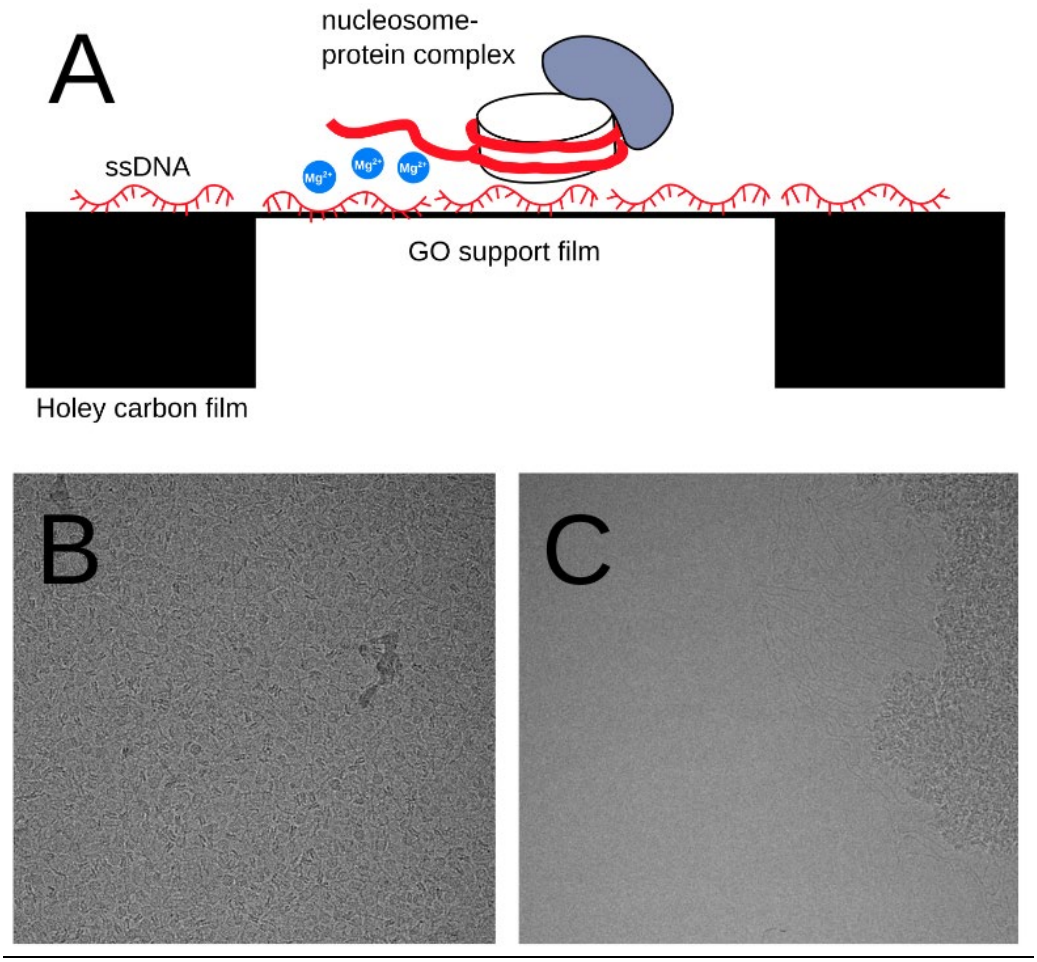

Figure 3.1: Nucleic acid grids stabilize chromatin specimens during vitrification. (a) Schematic of the proposed nucleic acid grids. Graphene oxide flakes span the holes in a holey carbon EM grid. Single-stranded oligonucleotides (red) are covalently bound to the GO surface. Chromatin complexes are expected to bind non-specifically through divalent cation-mediated ionic crosslinking of phosphate backbones on the ssDNA and double stranded nucleosomal DNA. (b) Cryo-EM image of snf2h-nucleosome complexes bound to the surface of a nucleic acid grid. (c) Cryo-EM image of a hole from the same grid as (b) but without a ssDNA-GO substrate. Long filaments are double stranded DNA.

We then imaged frozen-hydrated snf2h-nucleosome complexes bound to the nucleic acid grids. From previous experiences imaging this specimen, we expected that the ionic strength of the buffer (70mM KCl and 5mM $MgCl_2$) would strongly promote surface denaturation during vitrification. However, where there was a GO support film, numerous nucleosome-shaped particles were observed without obvious signs of nucleosome unspooling or protein denaturation (**Figure**

3.1). Incomplete coating of the GO grids offered us an internal control for the stabilizing effects of the ssDNA-GO substrate: in imaged holes that did not contain a GO support film, any observed specimens are likely adhered to the air-water interface, as we rinsed excess unbound specimen off of the grid before plunge-freezing. In these holes, we observed large aggregates of nucleosomes and long filaments of unspooled double-stranded nucleosomal DNA (**Figure 3.1**). These results demonstrate that nucleic acid grids bind and protect chromatin specimens from vitrification-induced surface denaturation at the air-water interface.

High-resolution structures of snf2h-nucleosome complexes without crosslinking

We collected a dataset of the nucleic acid grid-bound snf2h-nucleosome complex on a 300kV Titan Krios with a K3 direct electron detector. After removing a large number of images that did not contain the ssDNA-functionalized GO substrate, we were left with 1797 images. With image processing we obtained two distinct high-resolution reconstructions of the snf2h-nucleosome complex: one with singly-bound nucleosome and another with a doubly-bound nucleosome. These states are predicted from negative stain EM and fluorescence spectroscopy experiments (Racki et al, 2009, Leonard and Narlikar, 2015).

We determined the singly-bound reconstruction to a resolution of 3.1Å (**Figure 3.2**). This structure is essentially identical to the previously-determined snf2h-nucleosome (Armache et al. 2019) but has better resolution and higher quality density for the enzyme.

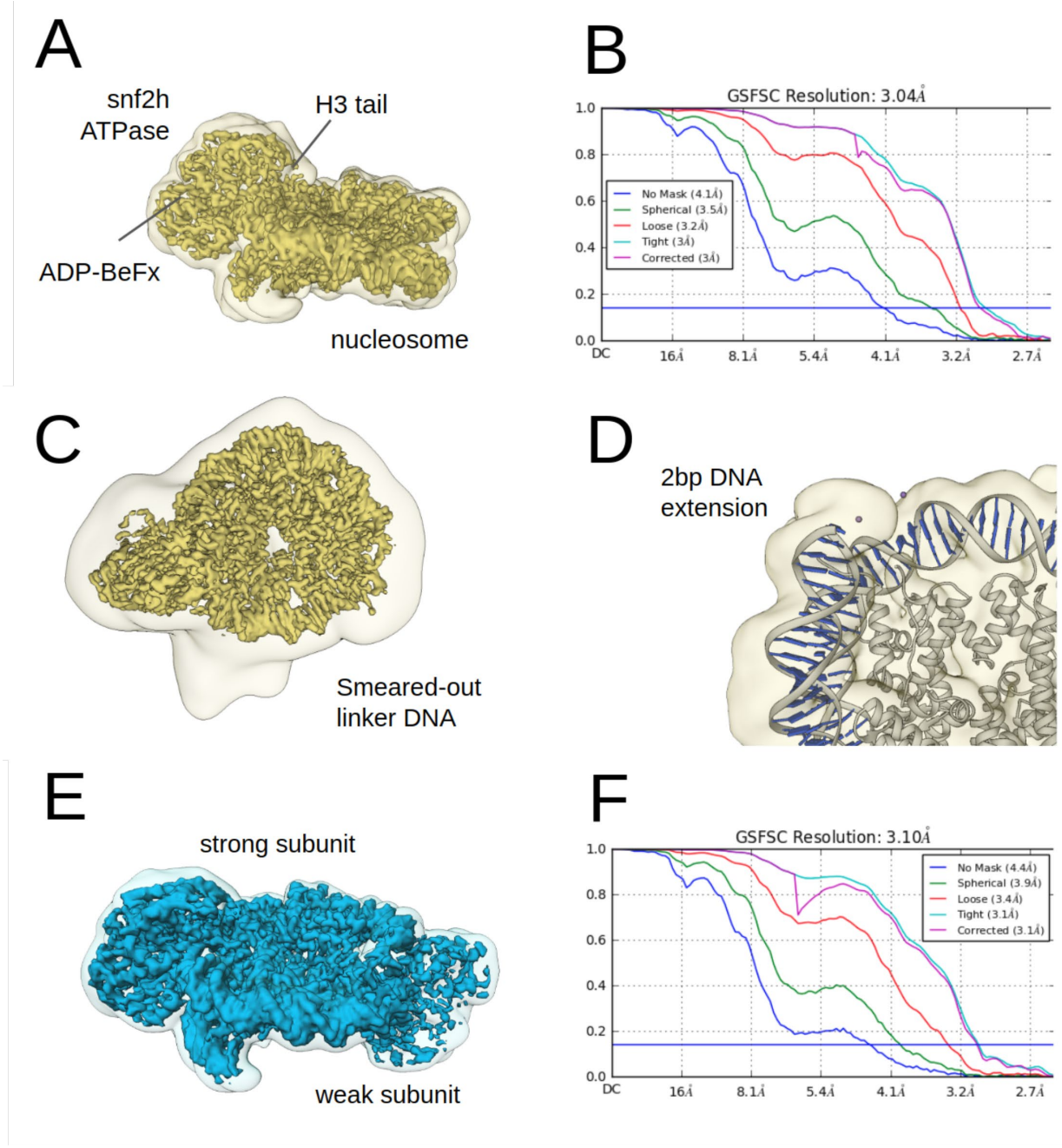

Figure 3.2: High-resolution reconstructions of uncrosslinked snf2h-nucleosome. (a) Density of the singly-bound snf2h-nucleosome complex at 3A resolution. (b) Gold-standard FSC curves for the snf2h-nucleosome complex with various density masks. (c) Top-view of the singly-bound complex. The transparent contour shows a low-isovalue of the density after low-pass filtering and demonstrates the location of the disordered linker DNA. (d) The density map shows a 2bp extension of nucleosomal DNA on the exit side. (e) Density of the doubly-bound snf2h-nucleosome complex at at 3.1A resolution. (f) Gold-standard FSC curves for the doubly-bound snf2h-nucleosome complex.

The snf2h ATPase domain appears to bind at superhelical locus 2 (SHL2) the the entry-side of 0-60 nucleosome substrate, as a smeared extension can be observed from the enzyme-bound side of the nucleosome at very low density isovalues (**Figure 3.2**). At the exit-side, an additional 2bp of dsDNA is observed compared to a nucleosomal core particle, consistent with the initial binding-induced DNA translocation observed in (Armache et al., 2019) (**Figure 3.2**). However, unlike previously determined structures of nucleosomes with linker DNA, the linker DNA is quite poorly resolved in our structures. This linker DNA may be bound by the linker DNA length-sensing hand-sant-slide (HSS) domain, which we do not resolve in our structures. We speculate that interactions of the linker DNA with the coarse surface of the ssDNA-GO substrate may lead to random bending or twisting, leading to incoherent averaging during 3D reconstruction.

Both the ADP and BeFx moieties are clearly resolved in the nucleotide binding pocket, confirming the state of the enzyme. Additionally, we see strong density for the basic patch of the H3 tail in the first recA lobe of the snf2h ATPase domain. Overall, this structure reproduces the results of Armache et al. (2019) but with higher-resolution structural data.

Through 3D classification and refinement, we also determined a doubly-bound structure of the snf2h-nucleosome complex. Snf2h has been proposed to work as a dimer to promote even spacing of the nucleosomes and our structure demonstrates how both catalytic subunits engage the nucleosomal substrate. The catalytic subunits are bound to the SHL2 and SHL-2 sites of the nucleosome, as expected from a low-resolution structure in Armache et al. (2019).

In this structure, both catalytic subunits engage the H3 tail and both apparently contain ADP-BeFx in the nucleotide pocket. However, density for one of the subunits is significantly weaker and more poorly resolved. While this could result from partial occupancy (singly-bound nucleosome particles contaminating the doubly-bound class), we think this is unlikely: extensive

3D classifications with a singly-bound template structure did not significantly improve this density, and low-pass filtering produces a density map with very similar enzyme densities which would not be the case if some particles did not contain a second enzyme. Instead, we believe the subunit with weak density is less tightly bound to the nucleosome. While we cannot determine from our structures what is causing one subunit to bind the nucleosome more tightly than the other, we speculate that it could have to do with asymmetric conformational changes to the nucleosomal substrate induced by the binding of the first promoter. We provide experimental evidence for such conformational changes in the next section.

Deformations in the histone octamer of a remodeler-bound nucleosome

In single-particle cryo-EM, the structure determination task is formally an inverse problem: given a forward model for how electron imaging of a 3D macromolecular structure produces noisy, aberrated 2D projection images, infer the 3D structure that generated the large set of experimentally-observed 2D projection images. If each 2D projection image was not generated from the same 3D macromolecular structure (i.e. if the macromolecular specimen is conformationally or compositionally heterogeneous), then a reconstruction determined by this approach will represent a weighted average of the true structures of all the particles.

If the average structure of a set of particles can be determined from 2D projection images, it may be possible to determine other information about the distribution of particle structures. For instance, the variance of density in a 3D reconstruction can be determined and corresponds to regions of conformational or compositional heterogeneity. Recently, it has also become possible to partially characterize the covariance of density in 3D reconstructions. In technical terms, the first few eigenvectors of the covariance matrix of the distribution of 3D density maps are estimated

from the particle images. Intuitively, these eigenvectors correspond to conformational 'reaction coordinates' and describe the most prevalent kinds of structural variability present in the specimen. By linearly interpolating a previously determined average structure along one of these eigenvectors/structural reaction-coordinates, it is possible to visualize continuous and discrete conformational changes that are not apparent from the mean 3D structure. This strategy is called '3D variability analysis' in the cryo-EM image processing software package cryoSPARC but has been proposed independently several times by other groups.

Unlike other approaches to characterizing conformational heterogeneity in single particle cryo-EM (like multi-body or normal-mode analysis), 3D variability analysis does not use the average structure to make assumptions about the nature of the conformational changes. However, because eigenvectors are linear directions of variability, complex large-magnitude nonlinear conformational changes cannot be represented with this approach. Additionally, the developers of cryoSPARC have not revealed exactly how their algorithm estimates these eigenvectors from the data. For this reason, while the results from 3D variability analysis are exciting and consistent with other biophysical data collected on the snf2h-nucleosome complex, we retain some skepticism about their precise interpretation.

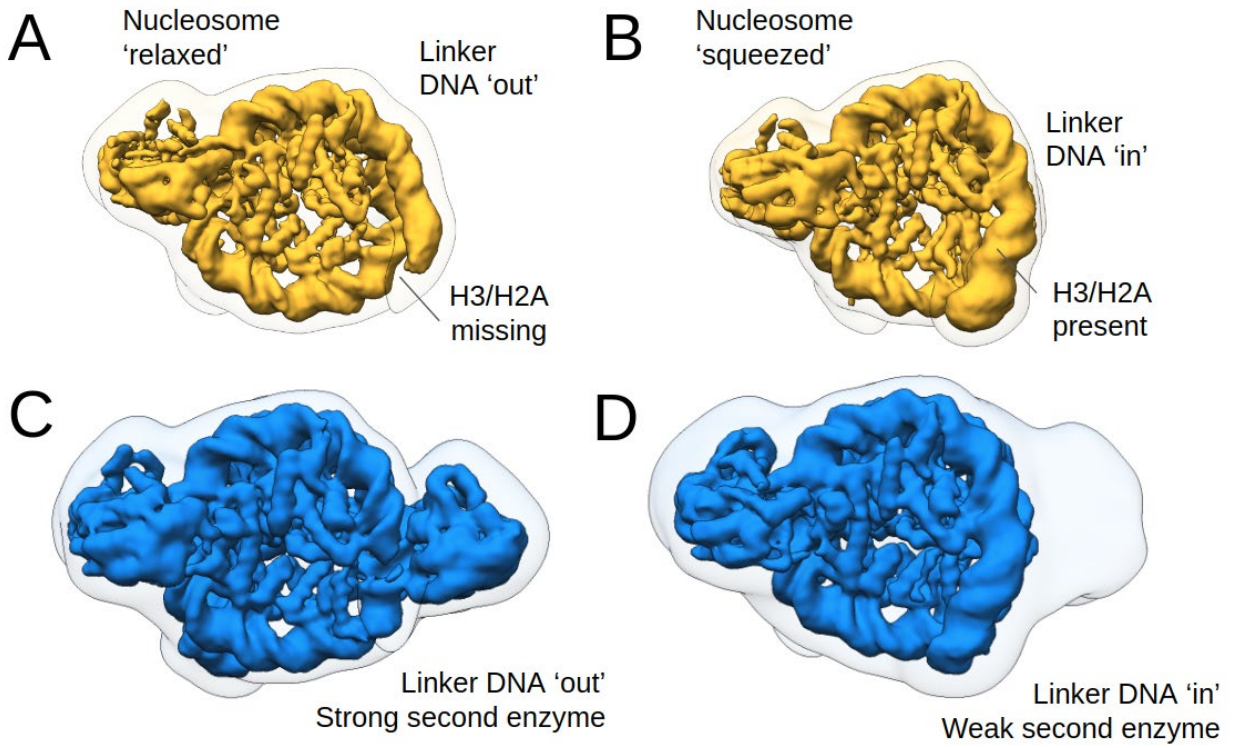

Figure 3.3: Deformations in the histone octamer of a remodeler-nucleosome complex. (a) Linker 'out' direction of the most significant conformational eigenvector calculated from the singly-bound complex. (b) Linker 'in' direction for the singly-bound complex. (c) Linker 'out' direction for most signficant conformational eigenvector calculated from the the doubly-bound complex. (d) Linker 'in' direction for the doubly-bound complex.

Keeping these concerns in mind, we applied 3D variability analysis to our singly- and doubly bound snf2h-nucleosome particles. In both subsets of particles, the most significant eigenvector showed a large breathing motion of the DNA near the nucleosome exit sites (**Figure 3.3**). In the positive direction of the eigenvector of the singly bound particles, DNA breathes outward, peeling away from the nucleosome core, and the entire nucleosome relaxes and widens (**Figure 3.3**). The nucleosomal DNA appears to become less tightly coiled around the octamer. Additionally, a helix of histone H3 and a loop from histone H2A disappears, significantly altering the local structure of the histone octamer. In the negative direction of the eigenvector, the linker DNA tightens around the histone octamer, squeezing into a more constrained, narrower

conformation (**Figure 3.3**). The H3 helix and H2A loop have strong density here. The snf2h enzyme appears more tightly associated with the DNA than the histone octamer, moving with the nucleosomal DNA as it becomes more or less tight.

In the first eigenvector for the doubly-bound particles, the major conformational changes associated (DNA breathing, octamer squeezing, histone deformation) are consistent with the first eigenvector from the singly-bound structure (**Figure 3.3**). Additionally, the density of the 'weak' snf2h subunit is strongly correlated with the linker DNA unpeeling from the octamer. When the linker DNA tightens around the nucleosome, the second snf2h subunit is only visible at very low density iso-values (**Figure 3.3**). This suggests (but does not prove) that conformational changes in the nucleosome, particularly the linker DNA may determine whether the nucleosome binding of second ATPase is tight or wobbly.

Discussion

We have developed nucleic acid grids that reliably stabilize fragile chromatin specimens through the harsh process of vitrification. We believe these grids work by binding chromatin specimens to an electron-transparent GO support film rendered anionic by functionalization with ssDNA. This promotes gentle, non-specific binding of chromatin by divalent cation crosslinking, holding the specimen away from the air-water interface formed during blotting and vitrification. These grids open up the broader class of chromatin specimen to being studied with cryo-EM in the absence of non-native stabilizing agents like chemical crosslinkers and in the presence of potentially destabilizing agents like monovalent salt and ATP that promote physiologically relevant conformational dynamics.

The structures we determined of snf2h-nucleosome complex are higher resolution than the previous structures determined in our lab. Nevertheless, they strongly support the conclusions drawn from that paper, including controversial claims that snf2h binding to the nucleosome causes a 2 base-pair translocation of DNA without multiple cycles of nucleotide turnover and that snf2h-bound nucleosomes have altered histone structure (Armache et al., 2019; Sinha et al., 2017). Considering the latter observation, novel 3D variability analysis of singly- and doubly-bound snf2h-nucleosome complexes support large structural changes in the nucleosome which seem to be driven by the overwinding and underwinding of the nucleosomal DNA and are coupled to the conformation of the H3 and H2A histone proteins. Further control experiments are required to test whether the conformational changes we observe in the nucleosome are induced and exacerbated by the binding of snf2h or if they are natural fluctuations in the structure of a nucleosome.

Materials and methods

Fabricating nucleic acid-coupled graphene oxide grids

Gold Quantifoil 200 mesh R1.2/1.3 holey carbon grids were coated with GO with according to our previously developed protocol (See chapter 2 and Palovcak et al. 2018). A 15-base single-stranded oligonucleotide (NH2-ssDNA) with an amino terminal group linked by a 6-carbon chain (purchased as a lyophilized powder from IDT, with sequence NH2-C6-GGTACCCGGGGATCG) was dissolved in DI water to a concentration of 20mM and stored at -20ºC. When preparing a batch of ssDNA-GO grids, the aqueous NH2-ssDNA was subsequently diluted in DMSO to 500uM. GO grids were submerged completely in 30uL of the DMSO/NH2-ssDNA solution and incubated at room temperature, shaking, overnight to functionalize. After

functionalization, ssDNA-GO grids were rinsed in 3 droplets of 100% EtOH followed by 3 droplets of DI water, blotted with Whatman 1 filter paper, and allowed to dry.

Preparation of the snf2h-nucleosome complex

Full-length wild-type snf2h was prepared as described in (Armache et al. 2019). Nucleosomes with 0bp of entry-side linker DNA and 60bp of exit-side linker DNA (0-60 nucleosomes) were prepared by salt dialysis from assembled octamers and double stranded DNA, essentially as described in (Armache et al. 2019). 100nM of snf2h was incubated with for 500nM of 0-60 nucleosomes in a buffer containing 2.0mM ADP-BeFx, 5.0mM $MgCl_2$, 70mM KCl, 20mM HEPES pH7.5, and 1.5% glycerol for 30 minutes at 4^{o}C.

Cryo-EM specimen preparation

3uL of snf2h-nucleosome complex was applied to a prepared ssDNA-GO grid and incubated for 1 minute. The grid was then blotted from the side onto Whatman 1 filter paper, rinsed in two 25uL droplets of snf2h-nucleosome buffer to wash off unbound complexes and side-blotted again. 3uL of additional buffer was applied to the specimen-bound grid to prevent drying and the grid was moved into a Mark IV vitrobot set to 100% humidity and 4^{o}C. The grid was then immediately blotted for 3 seconds with blot-force 0 before plunge freezing.

Cryo-EM data collection

Data was collected on a 300kV Titan Krios equipped with a K3 camera and a Gatan energy filter (GIF). The magnification was 105kx, giving a pixel size of 0.845Å/pixel. The electron dose was set to 0.7 electrons/pixel*frame. The total exposure was 50 electrons/pixel*exposure. Defocus

was set between 0.8um and 1.6um. Micrographs were collected using the multi-hole image shift protocol implemented in serialEM. 3174 movies were collected.

Cryo-EM image processing

Movie frames were aligned, summed, and electron dose-weighted using MotionCor2. Due to suboptimal coverage of the grid with GO, many images did not contain specimen-coated GO; these images were excluded by manual screening, leaving 1797 images with a thin coating of GO and visually-apparent snf2h-nucleosomes particles. CTFs were estimated with gCTF. Particles were picked and segmented using RELION (Kimanius, et al. 2016). All subsequent processing was performed with cryoSPARC2 (Punjani et al. 2017), including ab initio model generation, heterogeneous refinement, homogeneous refinement, and 3D variability analysis. All visualizations were performed using EMAN2 (images) and UCSF chimera (volumes and atomic models). All atomic models are from (Armache et al. 2019), docked into the new maps.

References

Armache, J.P., Gamarra, N., Johnson, S.L., Leonard, J.D., Wu, S., Narlikar, G., Cheng, Y. (2019). Cryo-EM structures of remodeler-nucleosome intermediates suggest allosteric control through the nucleosome. Elife.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife *5*.

Leonard, J., Narlikar G., (2015). A nucleotide-driven switch regulates flanking DNA length sensing by a dimeric chromatin remodeler. Mol. Cell. 57(5). 850-859.

Palovcak, E., Wang, F., Zheng, S., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A.,Cheng, Y. (2018). A simple and robust procedure for preparing graphene-oxide cryo-EM grids. J Struct Biol. *204*, 80-84.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296.

Racki, L., Yang, J., Naber, N., Partensky, P., Acevedo, A., Purcell, T., Cooke, R., Cheng, Y., Narlikar, G. (2009) The chromatin remodeler ACF acts as a dimeric motor to space nucleosomes. Nature. 462(7276). 1016-1021.

Sinha, K. K., Gross, J.D., Narlikar, G., (2017). Distortion of the histone octamer core promotes nucleosome mobilization by a chromatin remodeler. Science. 355(6322).

Chapter 4 - Generating contrast in cryo-EM images with convolutional networks

Contributing Authors

Eugene Palovcak[1], Daniel Asarnow[1], Zanlin Yu[1], Melody Campbell[1], Yifan Cheng[1,2]

[1]Department of Biochemistry and Biophysics, University of California San Francisco, CA 94143
[2]Howard Hughes Medical Institute, University of California San Francisco, San Francisco, CA 94143

Introduction

In single particle cryogenic electron microscopy (cryo-EM), high-resolution 3D structures of biological macromolecules are obtained by coherently averaging a large number of very noisy 2D projection images (Cheng 2015). This process requires identification of single particles in a brightfield image (particle picking) followed by sorting the particle images according to conformational state and viewing direction (classification and alignment). The success of each of these steps is critically dependent on the signal-to-noise ratio (SNR) of the cryo-EM images (Jensen 2001). The SNR of cryo-EM images is fundamentally limited by the radiation sensitivity and optical transparency of the frozen-hydrated cryo-EM specimen, which necessitates low-dose phase-contrast imaging (Glaeser 1999).

Within these constraints, the SNR of cryo-EM images is typically well below 1 (more noise than signal) and particles are barely visible in brightfield images (Frank and Al-Ali 1975). The conventional approach to maximizing SNR is to increase the defocus of the objective lens. Defocus alters the contrast transfer function (CTF) of the electron microscope and produces more contrast

at low spatial frequencies. However, high defocus also suppresses SNR at the intermediate and high-spatial frequencies required for accurate alignment and classification (Cheng 2015). In general, producing phase contrast with defocus forces the electron microscopist to balance particle visibility (SNR at low spatial frequencies) with the high-resolution information content of the image (SNR at high spatial frequencies). This issue becomes more acute for small or irregularly shaped particles and greatly increases the difficulty of determining the 3D structures of such specimens with cryo-EM (Herzik, Wu, and Lander 2019).

Ideally, there would be a way to improve contrast at low spatial frequencies without perturbing the information at high spatial frequencies. In optical microscopy, this is accomplished with quarter-wave phase plates, but such devices are just being invented for electron microscopes and have not found routine use in single particle cryo-EM (Danev, Tegunov, and Baumeister 2017; Schwartz et al. 2019). Complementary and alternative methods for generating contrast could therefore be of great value for single particle cryo-EM.

Here, we present a computational image restoration approach to enhance contrast and reduce noise in cryo-EM images. The basic idea is to train a parameterized image operator (a convolutional neural network or CNN) as an image denoiser. This training scheme, called noise2noise, uses noisy cryo-EM data itself as a training signal and is fully compatible with existing strategies for cryo-EM data acquisition (Lehtinen et al. 2018). We show that CNNs trained with noise2noise significantly enhance the contrast (low-frequency SNR) of cryo-EM images, similar to the effects of a quarter-wave phase plate. At high frequency, denoising CNNs also reduce noise but do so by transforming uncorrelated noise into a systematic false signal or 'bias'. While this bias makes denoised images unsuitable for calculating high-resolution reconstructions directly, they may still be useful in image processing steps, such as initial evaluation of

micrographs, particle picking, and initial alignment and classification steps. Furthermore, we show that high-frequency bias can be minimized by merging the low-resolution information from a denoised image with the high-resolution information from the unmodified original image, similar to the classical strategy of focal pair merging (Ludtke and Chiu 2003).

Training a convolutional neural network to denoise cryo-EM images

CNNs are powerful parameterized function approximators (Krizhevsky, Sutskever, and Hinton 2017). They consist of a large number of small convolution filters with learnable parameters. Each convolutional filter is applied to the input image in real-space and passed through a pixel-wise nonlinear function with a shape parameter (typically a threshold-based masking function called a 'rectified linear unit' or 'ReLU'). By stacking a large number of these simple operations in series (convolutional layers), complex input-output maps (functions) can be approximated (Goodfellow, Bengio, and Courville 2016).

Since each operation in a CNN is differentiable, the parameters of a CNN can be learned from a large set of input-output image pairs using gradient-based stochastic optimization. To train a CNN to approximate an image denoiser, these input-output pairs typically consist of an image with and without noise(Zhang et al. 2017). Training consists of calibrating the CNN's parameters such that applying the CNN to the noisy image produces the noiseless image. To ensure that the CNN generalizes to unseen images, the image content and the distribution of the noise in the training dataset must be representative of those expected during use.

These requirements pose a major problem for training a denoising CNN for cryo-EM images: the radiation sensitivity of frozen-hydrated biological material makes it impossible to observe a noiseless cryo-EM image in principle (Glaeser 1999).

Fortunately, learning the parameters of a denoising CNN does not actually require noiseless images. Lehtinen et al. recently demonstrated that pairs of noisy images of the same scene can be used in place of noisy/noiseless pairs to train denoising CNNs (Lehtinen et al. 2018). Training is performed as if the noisy image pair were a noisy/noiseless pair: the CNN is applied to the first noisy image and the discrepancy between the output and the second noisy image is calculated (the loss). The loss is used to calculate the direction each parameter should be updated to make the output and second noisy image more similar (the loss gradient). This parameter update is like the parameter update we would have calculated if the second image was noiseless, but perturbed in a random way by the noise in the second image. Lehtinen et al. showed that given a sufficiently large set of noisy image pairs, these perturbations average out and converge on the same parameter values that would be obtained using conventional noisy/noiseless training data.

This training strategy, termed noise2noise, enables the parameters of a denoising CNN to be learned without ever seeing noiseless images. It simply requires that the noise in the training pairs be statistically uncorrelated given the signal. In fact, noise2noise training makes no other assumptions about the structure of the signal or the distribution of the noise and learns both implicitly from the training data. A more mathematical description of the noise2noise training scheme is provided in the supplement and we also refer the reader to the original noise2noise paper and the more general framework for self-supervised denoising described in noise2self (Supplemental note 1, (Lehtinen et al. 2018; Batson and Royer 2019).

Because modern electron detectors collect cryo-EM images as dose-fractionated movies, we can generate two images with the same signal but uncorrelated noise simply by summing up disjoint sets of movie frames, such as the even and odd frames after motion correction and electron dose-weighting (Zheng et al. 2017; Grant and Grigorieff 2015). The noise2noise scheme therefore requires no alteration in the data collection scheme and can work on previously collected datasets.

Implementing a denoising CNN for cryo-EM

In principle, the noise2noise training scheme only requires a large dataset of noisy image pairs to train a denoising CNN. In practice, we found that denoising performance depended strongly on the preprocessing of the training images and the structure of the CNN.

We train one CNN per dataset and denoise the same images that we were used to train the CNN. This is possible because a typical cryo-EM dataset consists of thousands of movies and the training process is relatively fast (a few hours). With such an abundance of training data, overfitting the CNN is very unlikely. Additionally, if the parameters of the CNN had been overfit, this would imply the CNN had memorized which pairs of noisy images went together. Overfitting would then be easy to detect: the output images would not be denoised.

To generate training data from a set of dose-fractionated electron movies, we first generate motion-corrected and dose-weighted sums of the even and odd frames using our program MotionCor2 (Zheng et al. 2017). Each of these images can range from 16 to 24 megapixels with a pixel size of approximately 1Å.

Each cryo-EM image is distorted by the contrast transfer function (CTF) of the electron microscope. Importantly, the CTF of each image is different and will distort otherwise similar signals in unique but predictable ways. A denoising CNN would need to learn to distinguish all

CTF-aberrated signals from noise, which is a more challenging task than simply learning to distinguish all unaberrated signals from noise. We chose to correct the CTF in the noisy images upfront: we estimate the CTF with the program gCTF and then multiply the Fourier transform of the image by the sign of the CTF and compute the inverse Fourier transform (phase-flipping) (Zhang 2016).

We reasoned that information at very high spatial frequencies is almost entirely noise (Baxter et al. 2009). Since we do not expect any denoising algorithm to recover the signal under such conditions, we chose to downsample training images. We downsample by cropping the Fourier transform of an image and computing the inverse Fourier transform (Fourier cropping). This has the effect of reducing the total noise in the image and making it smaller. We typically Fourier crop the images so their pixel size is 2Å (Nyquist frequency=1/4Å).

To facilitate training in batches, we break up the phase-flipped, Fourier-cropped training images into square patches (typically 192x192 pixels). For most macromolecular specimens examined, a patch of this size contains several particles. We normalize each patch by subtracting the mean pixel value and dividing by the standard deviation.

We used a CNN architecture similar to the U-net used by Lehtinen et al. with some minor modifications (Ronneberger, Fischer, and Brox 2015). First, we replaced each block of convolutional layers in the U-net with a wide-activation convolutional layer, which has recently been described in state-of-the-art super resolution models (Yu et al. 2018). Second, we used depth-to-space upsampling to minimize aliasing artifacts (Odena, Dumoulin, and Olah 2016). We found these modifications improved the training loss and the visual quality of the output compared to a standard U-net. In the supplement, we provide a full description of the CNN architecture used in this work (**Figure S4.1**).

Finally, we train the CNN using the Adam stochastic optimization algorithm with weight normalization (Kingma and Ba 2014; Salimans and Kingma 2016). We find that training CNNs with noise2noise is fast, typically achieving a stable plateau in the value of the loss function in several epochs (20-30 minutes on a single modern GPU). We train for one hundred epochs.

At test time, we preprocess images in the same way we did during training, except we do not break the image into patches. This is possible because CNNs can operate on images with arbitrary shape and size. Finally, we upsample the denoised image by zero-padding its Fourier transform, low-pass filtering with a soft radial Fourier mask, and calculating the inverse Fourier transform. The result is a denoised image with the shape and pixel size of the original image, bandlimited to the same frequency as the training data.

All code was written in python. Neural networks were implemented and trained using the libraries keras and Tensorflow (Abadi et al. 2016).

Contrast enhancement for diverse cryo-EM specimens

We trained denoising CNNs on several previously collected cryo-EM datasets with a range of specimen types and imaging conditions: the Falciparum 80S ribosome (~3.2 megadaltons), the human TRPM4 ion channel (~700 kilodaltons), a human integrin-Fab complex (~380 kilodaltons), and the human PKA catalytic domain (~43 kilodaltons). Despite differences in particle shape, molecular weight, and imaging conditions, all denoised images show significantly enhanced contrast (**Figure 4.1**). Particle shapes are clearly defined and broadly consistent with previously-determined 3D reconstructions (Wong and Scheres 2019; Autzen et al. 2018; Cormier et al. 2018; Herzik, Wu, and Lander 2019).

To begin to understand how a denoising CNN effects information at different spatial frequencies in the image, we trained a denoising CNN on a dataset of a well-characterized test specimen, the Thermoplasma acidophilum 20S proteasome (T20S) (**Figure 4.2**). In denoised images of the T20S, Fourier amplitudes are significantly higher at low spatial frequencies than in the original image (**Figure 4.2**). Thon rings associated with signal-dependent modulation by the CTF are relatively weaker but still present and correctly located in denoised images: defocus values estimated from denoised images are mostly quite close to those estimated from the original noisy image (**Figure 4.2**). The small set of images that have very different estimated defocus values after denoising are heavily contaminated with crystalline ice.

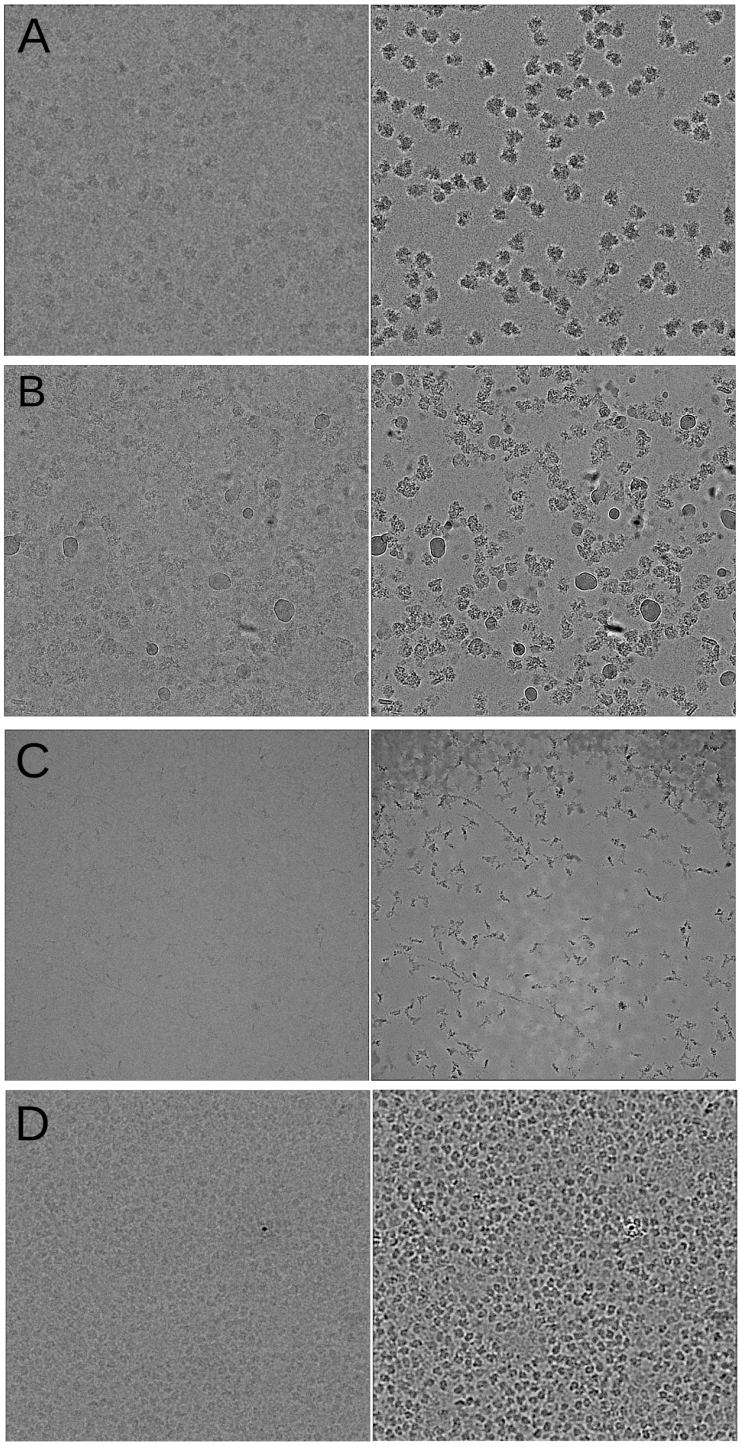

Figure 4.1: Performance of denoising CNNs on real cryo-EM images. (a) Falciparum ribosome particles, (b) TRPM4 particles, (c) integrin-Fab particles, (d) protein kinase A particles, before (left) and after (right) denoising.

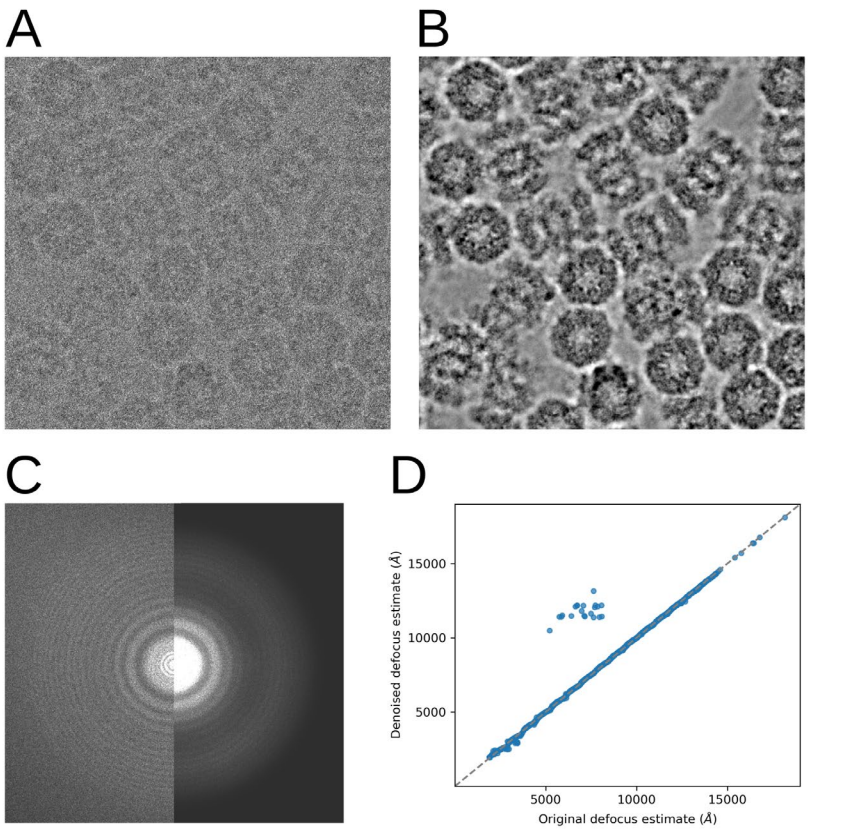

Figure 4.2: Effects of denoising on Fourier amplitudes. 20S proteasome particles before (a) and after (b) denoising. (c) Amplitude spectrum of the 20S proteasome particles before (left) and after (right) denoising. (d) Scatter plot of average defocus values estimated from images before and after denoising. Defocus values were estimated using gCTF and the major and minor defocus values were averaged. The small population of off-diagonal images (24 of 843) appear heavily contaminated with crystalline ice.

This demonstrates that CNNs trained by the noise2noise scheme are effective at denoising and contrast enhancement on a wide range of real single-particle cryo-EM specimens. Denoising occurs primarily by boosting the Fourier amplitudes at low spatial frequencies but does not completely suppress Fourier components at high spatial frequency. We show four cases here, but we have not found any case where the contrast of single-particle cryo-EM specimens is not significantly enhanced by training and applying a denoising CNN.

Quantitative signal enhancement and noise suppression

What is the denoiser doing to the signal and noise in the image? How much should we trust the contrast enhancement we see in the denoised images? Is it conceivable that denoised images could be used in single-particle reconstructions?

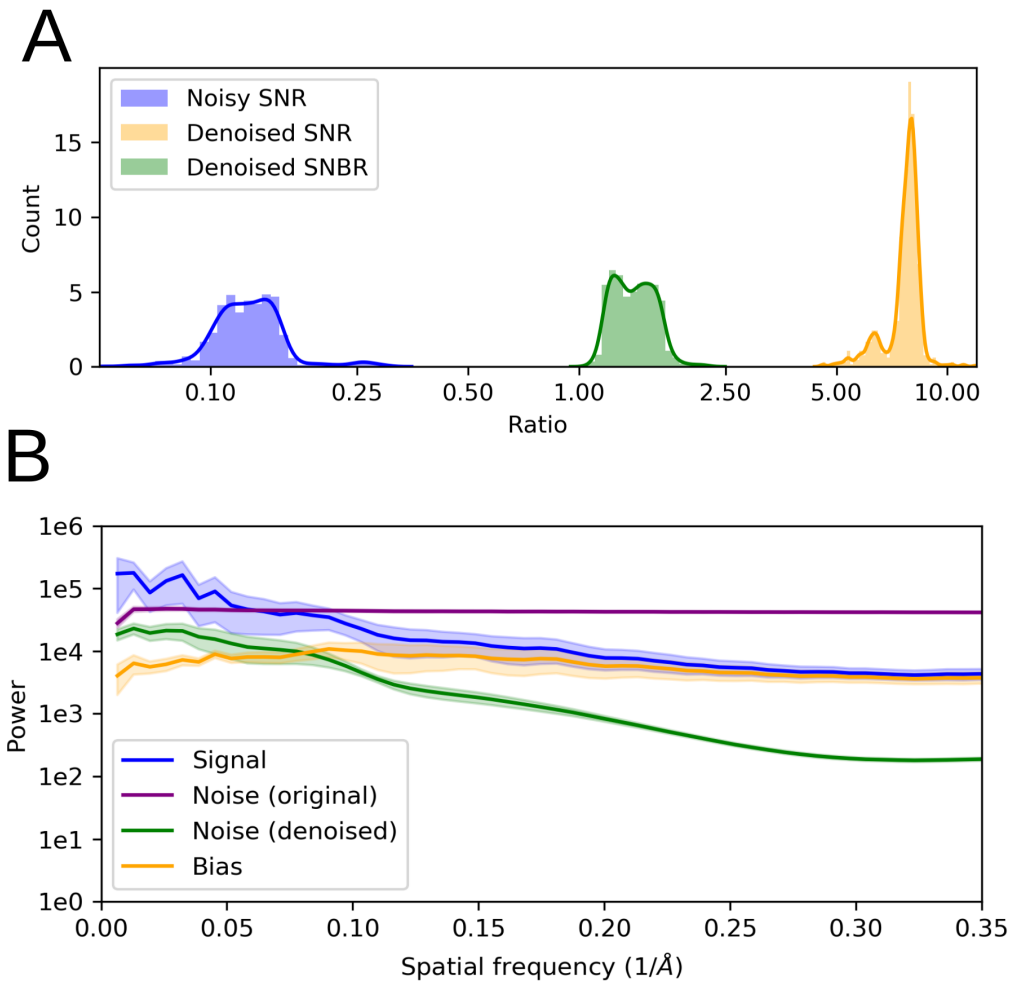

Figure 4.3: Quantitative analysis of signal, noise, and bias in denoised images. (a) Histograms of SNRs before (blue) and after (orange) denoising and SNBRs after denoising (green) for images from the T20S proteasome dataset. Smooth lines represent kernel density estimates of the distribution. The x-axis is on a logarithmic scale. (b) Spatial frequency-dependent variance (power) of the signal (blue), bias (orange), and

noise. Noise power is calculated before (purple) and after (green) denoising. The y-axis is on a logarithmic scale. Curves represent the mean of the quantities for all images in the T20S proteasome dataset. Shaded regions show one standard deviation above and below each mean curve. All quantities were calculated as described in Supplemental notes 2 and 3.

To answer these questions, we extended the conventional cross-correlation-based estimators for SNR and spectral SNR (SSNR) to handle images modified by a deterministic, arbitrary operation such denoising (Supplemental note 2, (Frank and Al-Ali 1975; Bershad and Rockmore 1974; Baxter et al. 2009). Our approach estimates the magnitude of any 'false signal' or 'bias' added to each image during denoising. This is possible because pairs of denoised images of the same scene will share signal and bias in common, while a denoised image and its paired noisy image will only share the signal. By estimating the signal variance, bias variance, and noise variance, we can compute quantities like the the signal-to-noise ratio (SNR) and similar quantities that take bias into account. Importantly, we can also compute these quantities as a function of spatial frequency.

We estimated the SNR for noisy and denoised images of the T20S dataset (**Figure 4.3**). While the mean SNR of the noisy images is ~0.15, the mean SNR of the denoised images is ~7.5, a 50-fold improvement. This improvement could have resulted from two distinct possibilities. First, the noise could have been removed, increasing the relative power of the signal. Second, the noise could have been transformed into bias that is correlated between denoised images but is not correlated with the true signal. We define another quantity, the signal-to-noise-and-bias ratio (SNBR), which is the ratio of the signal variance and the sum of the noise and bias variances (Supplemental note 2). Intuitively, the SNBR represents the relative strength of true signal compared with the strength of all other components in the image. For the T20S dataset, the mean SNBR is ~1.5, a 10-fold improvement over the original SNR of the noisy images.

We also estimated the frequency-dependent variance (power) of the signal, bias, and noise before and after denoising (**Figure 4.3**). When we plot the average of these quantities over all micrographs in the T20S dataset, we find that the noise in the original image dominates the signal in all but the lowest spatial frequencies. After denoising, the noise is much smaller than the signal at all spatial frequencies. Additionally, the signal is significantly larger than the bias at low spatial frequencies (>0.1 1/Å) but has similar power at higher spatial frequencies.

Taken together, these results indicate that denoising increases the strength of the true signal in the denoised images by a sizeable factor. However, the denoising also transforms uncorrelated noise into statistically correlated bias, which is undesirable and is almost as strong as the signal at intermediate and high spatial frequencies. What might be the nature of this bias? One possibility is that in highly under sampled cryo-EM images, very high spatial frequency patterns in the signal may be impossible to disambiguate. For example, the precise arrangement of side chains in an image of a folded protein will be encoded by very high-spatial frequency Fourier components. When heavily corrupted with noise, an image of such a signal may have several plausible interpretations. Because the CNN could be wrong about any particular arrangement of side chains, the best guess for minimizing the mean-squared error is a pixel-wise average over all possible interpretations. This average would appear blurred and match no single arrangement of side-chains exactly, but would not be far off (in the mean-squared error sense) from any single plausible arrangement.

If this is the case, low spatial frequencies should be relatively less biased than high spatial frequencies in denoised images, because the higher SNR at low spatial frequencies should make it easier for the CNN to identify the signal unambiguously. This is consistent with the denoising

CNN enhancing the true signal at low spatial frequencies while mostly transforming the noise into bias at high spatial frequencies.

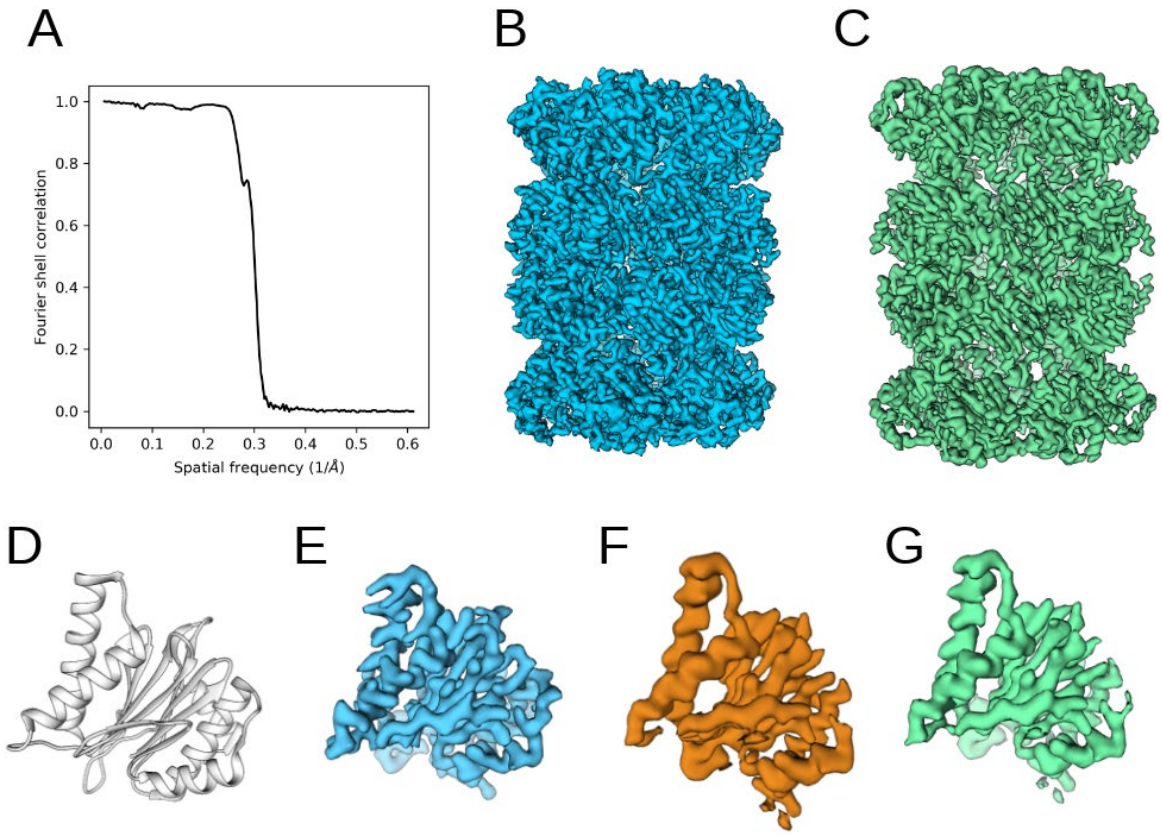

Figure 4.4: 3D reconstructions of denoised 20S proteasome particles. (a) FSC calculated comparing the 20S proteasome maps calculated from the noisy particles (b) and the denoised particles (c). (d) Ribbon model for one beta subunit of the 20S proteasome. Densities for the beta subunit of the 20S proteasome map calculated from back-projecting the noisy particles (e), denoised particles without sharpening (f), and the denoised particles with a -40Å^2 b-factor.

3D reconstructions with denoised particles

Even though the signal power is strongest at low-spatial frequencies in denoised images, our measurements show that it is finite at all spatial frequencies up to the band-limit of the training data. This suggests that denoised images might be useful for image processing tasks that require image details to be accurate at intermediate and high spatial-frequencies. To test this, we back-

projected denoised 20S proteasome particles using the alignments derived from refining the noisy particles in cryoSPARC. We calculated the Fourier shell correlation (FSC) between the map from the noisy particles and the denoised particles. If the denoising algorithms significantly distorted the intermediate and high spatial frequency information in the particle images, we would expect this FSC to become decorrelated at these frequencies.

Instead, we see that the denoised reconstruction is highly correlated (FSC>0.9) with the noisy reconstruction up to 1/4Å, where it begins dropping off (**Figure 4.4**). The FSC curve is near zero at 1/3Å. Additionally, the maps are visually similar, especially in the overall shape and subunit topology of the 20S proteasome (**Figure 4.4**). This is expected, because the denoiser was trained on and applied to images downsampled to 1.5Å/pixel. We upsampled the denoised images to a pixel size of 0.814Å, but the image is bandlimited in Fourier space beyond 3Å.

Comparing the maps in the well-resolved beta subunit, the denoised reconstruction has clearly produced overall the correct structure, although the low-resolution features of the map seem over-weighted with respect to the map calculated from noisy particles (**Figure 4.4**). This might be expected, given that a major effect of the denoiser is to boost the relative amplitudes of low-frequency Fourier components in the images. When we apply a -40Å2 B-factor to the denoised map, some of the high-resolution features (such as side-chains) are recovered (**Figure 4.4**).

These reconstructions are highly consistent with the estimates of the signal variance from the previous section: signal is largely preserved in the denoised images and averages to produce a reasonably correct reconstruction up to the bandlimit of the denoiser's training data. However, the reconstruction is not perfect, as even very low spatial frequencies in the FSC curve are highly but not perfectly correlated (**Figure 4.4**). Additionally, the amplitudes of the denoised images are not

correct and high-resolution details such as side chains only appear in the denoised maps when a significant negative B-factor is applied (**Figure 4.4**).

Conclusions

We have demonstrated that trained denoising CNNs can produce significant low-resolution contrast in cryo-EM images. Training with the noise2noise scheme obviates the need for clean training data and is possible with previously collected cryo-EM images. Visually, denoised images resemble images taken with a Volta phase plate and can be used to judge the quality of cryo-EM specimen before committing to large-scale automated data collection. Denoised images will also likely be useful for particle picking, especially for small or oblong particles (**Figure 4.1**).

For quantitative imaging sciences like cryo-EM, it is insufficient for image restoration algorithms to simply make the image look better. The signal must be enhanced relative to the noise and any false signals (bias) must be kept to a bare minimum. We developed a quantitative frequency-dependent framework for measuring signal, noise, and bias in denoised images. For our trained denoising CNNs, we found that while the signal-to-noise ratio is indeed significantly enhanced after denoising, most of the noise suppression at high spatial frequencies results from the conversion of noise into bias. At lower and intermediate spatial frequencies (up to 1/10Å in the case of the T20S proteasome), the signal power is higher than the bias power and both are significantly higher than the noise power. Additionally, reconstructions performed with denoised particles are largely similar to reconstructions with the original noisy particles, suggesting that even if the bias is present in the images, it largely averages out when particles are back-projected.

We have contributed a computer program, *restore*, for training denoising CNNs and applying them to real cryo-EM datasets. We have also introduced the metrics required to

understand how denoising CNNs effect images. While we were preparing this manuscript, we learned of several other groups using denoising CNNs with the noise2noise training scheme in similar ways, achieving similar results. It will be interesting to use our quantitative metrics to determine which denoising CNNs perform the best.

We have shown that denoised images contain information at high-spatial frequencies and may be useful for tasks like classification and alignment. These applications could be highly impactful for small or highly flexible particles. Nevertheless, there are still hurdles to using denoised images in cryo-EM, as the presence of bias in the images invalidates the assumptions of the gold-standard FSC. Even so, we are optimistic: unlike contrast enhancement methods like the VPP, denoising is computational and therefore reversible. It is also straightforward to merge denoised information at low spatial frequencies with noisy, unperturbed information at higher spatial frequencies to minimize bias. Determining appropriate ways of leveraging the enhanced signal in denoised images for classification and refinement are important future directions.

Supplemental note 1: Mathematical description of the noise2noise training scheme

First, we assume an image (M) to be a vector of pixels composed of a signal vector (S) and an additive noise vector (N). Though we do not explicitly denote it, the noise component may or may not depend statistically on the signal and therefore can include shot noise (Poisson noise).

eq.1 $$M = S + N$$

A parameterized denoiser is a function with parameters θ that takes the noisy image, M, and outputs the signal, S:

eq.2 $$f_\theta(M) = S$$

The conventional strategy for training a deep CNN to be an image denoiser is to search for the set of parameters, θ, that minimizes the expected value of a loss function (the expected risk) over all the clean/noisy image pairs in the training data:

eq.3 $$\text{argmin}_\theta \, E(|f_\theta(M) - S|^p) = \text{argmin}_\theta \, E(|f_\theta(S+N) - S|^p)$$

where $E(X)$ denotes expectation over the data distribution and $|X|^p$ denotes the L_p norm over all pixels in the resulting difference image. In this and other related works, $p=2$ and corresponds to the mean-squared error over pairs of corresponding pixels, though L_1 (median-seeking) and approximate L_0 (mode-seeking) norms were also demonstrated in Lehtinen et al. These parameters are determined using some variant of stochastic gradient descent over many small batches of the training data. Because f_θ (and any CNN in general) is a differentiable function, the gradient of each parameter with respect to the loss can be estimated directly using the chain rule (the back-propagation algorithm).

The key insight of Lehtinen et al. is that one can learn the same parameters without clean data. Instead of M and S, assume we have access to two images with the same signal but uncorrelated noise:

eq.4 $$M_1 = S + N_1$$

$$M_2 = S + N_2$$

We make two very general assumptions about the noise:

eq.5 $$P(N_2|S,N_1) = P(N_2|S)$$

eq.6 $$E(N_i) = \mathbf{0}$$

where $\mathbf{0}$ is a vector of zeros with the same shape as the images M_1 and M_2. Then, we train the CNN with stochastic gradient descent or one of its variants to minimize the following objective:

eq.7 $$\text{argmin}_\theta E(|f_\theta(M_1) - M_2|^p) = \text{argmin}_\theta E(|f_\theta(S+N_1) - S - N_2|^p)$$

In words, we are training f_θ to convert one noisy image, M_1, into the second noisy image M_2. Clearly this is not what we actually want f_θ to do. However, because of (eq.5), M_1 provides f_θ with *no information* about the specific instantiation of noise N_2 that it is tasked to predict: N_2 is an independent random draw from the noise distribution $P(N|S)$. To minimize the expected risk, the best guess f_θ can make for N_2 is the average of all the possible instantiations of noise it might see, which by (eq.6) is zero noise. This implies that the objective in (eq.7) is minimized by the same set of parameters $\boldsymbol{\theta}$ that minimizes (eq.3), and so (eq.7) trains f_θ to be a denoiser.

The *noise2noise* strategy is more intuitive when considered geometrically. Recall that an image is a vector of pixel intensities, locating a single point in a high-dimensional vector space. For an image with two pixels, that space would be a plane. A signal, S, occupies a point in the space. Noise is a vector N_i that displaces a noisy image $S+N_i$ to a point away from S. The

conventional denoising CNN strategy trains a mapping between a given noisy point $S+N_i$ and S. The *noise2noise* strategy attempts to train a mapping between a given noisy point $S+N_1$ and a second noisy point $S+N_2$, but provides no information about N_2. Because N_2 could be any of the possible noisy points and we penalize f_θ if its output is distant from $S+N_2$, the least risky guess is the point that is closest to all of the noisy points. This point is simply the mean of all noisy images $S+N_i$, which is also the noiseless signal, S.

Supplemental note 2: Estimating signal, noise, and bias in denoised cryo-EM images

Once we have denoised cryo-EM images with a trained CNN, we would like to evaluate the SNR of the denoised image and the magnitude of any systematic errors added by the denoiser. Additionally, we would like to estimate these quantities as a function of spatial frequency. We assume that a denoised image (D) is the sum of the signal (S), the remaining uncorrected noise (N_d), and some false signal (B) that results from systematic error in the denoiser (the bias of the estimator). We distinguish N_d from the noise in the original image, which we refer to as N_n.

eq.8 $$D = S + N_d + B$$

The signal-to-noise ratio is defined as the ratio of the signal variance and the noise variance:

eq.9 $$SNR = var(S)/var(N)$$

Thus, we need to estimate var(S), var(N_d), and var(N_n). We also would like to estimate var(B) for the denoised images. We recall several elementary identities of the variance and covariance:

eq.10 $$\mathrm{var}(X) = \mathrm{cov}(X,X)$$

eq.11 $$\mathrm{cov}(X,Y) = E([X-E(X)][Y-E(Y)]) = E(XY) - E(X)E(Y)$$

eq.12 $$\mathrm{cov}(U+V, Y+Z) = \mathrm{cov}(U,Y) + \mathrm{cov}(V,Y) + \mathrm{cov}(U,Z) + \mathrm{cov}(V,Z)$$

eq.13 $$\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X,Y)$$

eq.13 is implied by eq.10 and. eq12. If X and Y are statistically independent, E(XY)=E(X)E(Y) and eq.11 implies that cov(X,Y)=0.

Computationally, the covariance of two images can be estimated with:

eq.14 $$\mathrm{cov}(X,Y) = N^{-1} \Sigma_i (X_i-\bar{X})(Y_i-\bar{Y})$$

Here, N is the number of pixels in images X and Y, X_i is the intensity of the i^{th} pixel of image X, and $\bar{X}$ is the mean pixel intensity.

We denote the noisy even and odd image sums as M_1 and M_2 and their denoised versions as D_1 and D_2. From the previous equations, it is straightforward to show with arithmetic that:

eq.15 $$\mathrm{var}(S) = \mathrm{cov}(M_1,M_2)$$

eq.16 $$\mathrm{var}(N_n) = [\mathrm{cov}(M_1,M_1)\mathrm{cov}(M_2,M_2)]^{1/2} - \mathrm{cov}(M_1,M_2)$$

eq.17 $\text{var}(N_d) = [\text{cov}(D_1,D_1)\text{cov}(D_2,D_2)]^{½} - \text{cov}(D_1,D_2)$

eq.18 $\text{var}(B) = \text{cov}(D_1,D_2) + [\text{cov}(M_1,M_1)\text{cov}(M_2,M_2)]^{½} - 2\text{cov}(M_1,D_2)$

In eq.16, the first term on the right-hand side is var(S+N), estimated as the geometric mean of $\text{var}(M_1)$ and $\text{var}(M_2)$. Similarly, the first term on the right-hand side of eq.16 is var(S+N+B) and is estimated as the geometric mean of $\text{var}(D_1)$ and $\text{var}(D_2)$. From these equations, we can estimate the SNR of the images before and after denoising:

eq.19 $\text{SNR\_noisy} = \text{var}(S) / \text{var}(N_n)$

eq.20 $\text{SNR\_denoised} = \text{var}(S) / \text{var}(N_d)$

Intuitively, eq.19 is identical to the expression for the SNR of noisy image pairs used by the Frank group and originally suggested by Bershad and Rockmore (Frank and Al-Ali 1975; Bershad and Rockmore 1974):

eq.21 $\text{cc}(M_1,M_2) = \text{cov}(M_1,M_2)/[\text{var}(M_1)\text{var}(M_2)]^{½}$

eq.22 $\text{SNR\_noisy} = \text{cc}(M_1,M_2) / [1 - \text{cc}(M_1,M_2)]$

We can also estimate other potentially interesting quantities such as the ratio of the signal variance and the bias variance (signal-to-bias ratio, SBR).

eq.23 $\text{SBR} = \text{var}(S)/\text{var}(B)$

Finally, we can estimate each of these quantities as a function of spatial frequency. To estimate the covariance of X and Y at some spatial frequency k, we calculate:

eq.24 $$\mathrm{cov}(X(k), Y(k)) = N_k^{-1} \Sigma_i X(k)_i Y(k)_i^*$$

where X(k) is the ring of Fourier components in the Fourier transform of X with spatial frequency k, N^k is the number of Fourier components in ring k, Y(k)* denotes the complex conjugate of Y(k), and the sum runs over each corresponding pair of Fourier components, i. This approach treats each ring in the Fourier transforms of X and Y as an independent random variable.

Supplemental note 3: Technical description of the training and denoising implementation

Code implementing these methods is available at https://github.com/eugenepalovcak/restore.

I/O and image pre-processing

Our program, restore, takes as input a RELION-style STAR file containing the file path and CTF parameters for a set of training micrographs. Each image in this STAR file is assumed to be the sum of all frames summed together after motion correction and dose-weighting with MotionCor2 and CTF estimation with CTFFIND, Gctf, or some other CTF-estimation program. Each training image is expected to have a training pair of even/odd images. These can be generated with recent versions of MotionCor2 using the '-SplitSum 1' option. In this case, the full sum image has the suffix 'DW' while the even/odd half sum images have the suffixes 'EVN' and 'ODD'.

restore begins by preprocessing the training data. For each even and odd half-sum image, we load the MRC file into an array. Then, we Fourier crop the micrograph such that the pixel size is ½ the value of the '--max_resolution parameter', effectively setting the nyquist frequency '--max_resolution'. A band-pass filter is applied in Fourier space to remove low frequency information (lower than 1/200 angstroms, corresponding to gradients of image contrast due to changes in ice thickness) and high-frequency information beyond the Nyquist frequency.

Next, we correct for the CTF by phase-flipping. Explicitly, we calculate the amplitude modulation implied by the CTF for each Fourier component, compute the sign of each amplitude modulation (+1 or -1), and multiply the Fourier transform of the Fourier-cropped image by this array. We inverse Fourier transform this image.

Then, we break each even and odd half-sum image up into 192x192-pixel patches. We normalize by subtracting the mean and dividing by the standard deviation. We store the patches in a large hierarchical data format (HDF) file. This allows fast access to the preprocessed data during training without requiring the entire dataset to be loaded into memory.

CNN architecture

We specified the architecture of the CNN using the keras library. Our CNN has a U-net architecture with specialized CNN block structures. The encoder branch consists of an initial 2D convolutional layer followed by three larger blocks of convolutional layers that progressively down-sample the input. In the decoder branch, three blocks of convolutional layers then progressively upsample the feature maps and receive 'skip connections' from the encoder.

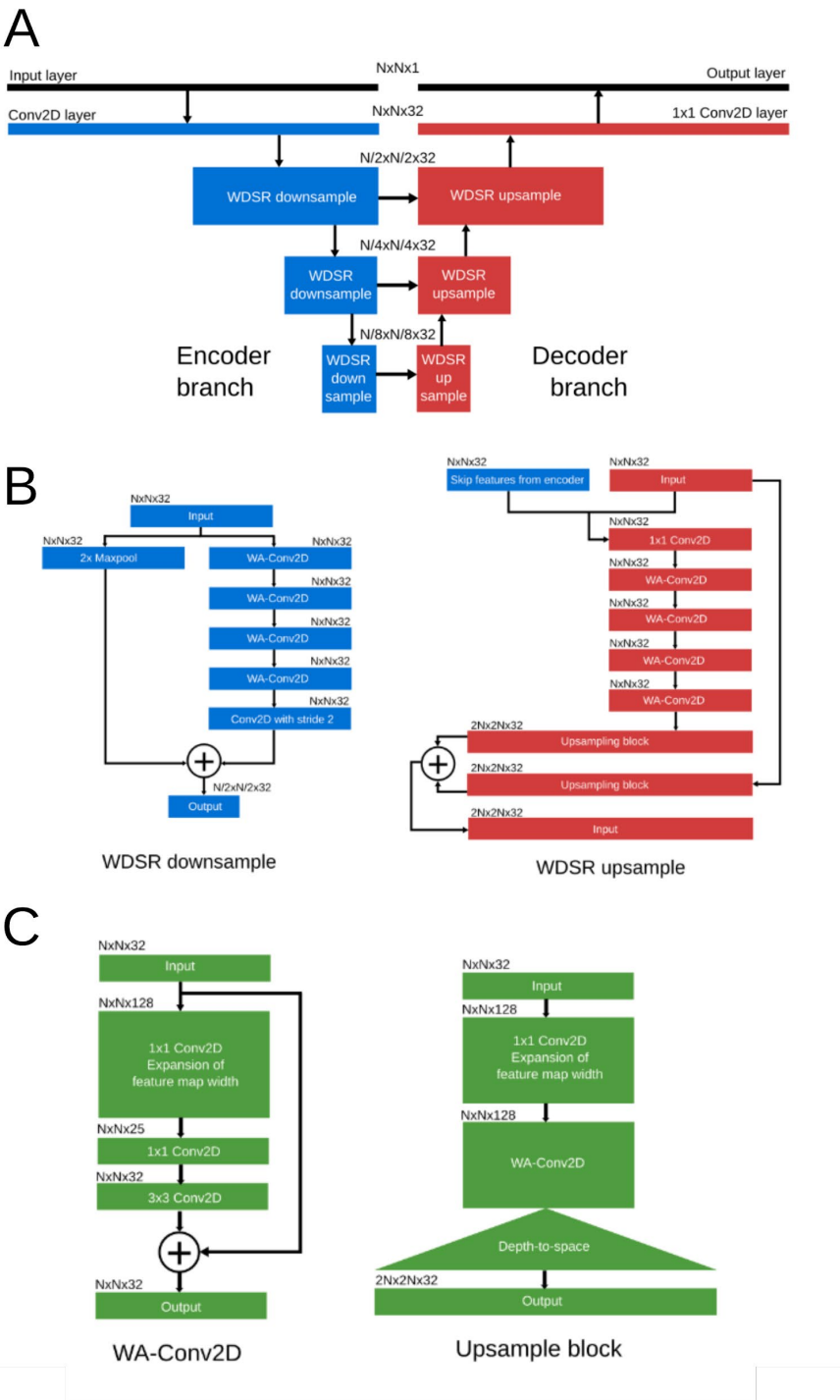

Figure S4.1: 3D reconstructions of denoised 20S proteasome particles. (a) WAUnet architecture, similar to a U-net-like architecture but with miniature WDSR networks for each of the convolutional blocks. (b) Miniature WDSR networks for downsampling and upsampling. (c) Layer schematic for the WA-Conv2D and upsample blocks that compose the WDSR networks.

The blocks of convolutional layers are modeled after the wide-activation super resolution (WDSR) networks described by Yu et al. The basic building block of these networks is the wide-activation convolutional layer. Wide activation convolutional layers take linear combinations of feature maps to increase their depth before applying a non-linearity and convolution operation. These expanded linear combinations allow more information to pass through the layers without getting decimated by the ReLU non-linearity (the so-called 'vanishing gradient' problem). For the same reason, each layer also uses residual connections, adding the processed information to the input.

Wide-activation convolutions are implemented by performing a 1x1 convolution to increase the depth of the input feature maps from 32 to 128, applying a ReLu non-linearity, and performing another 1x1 convolution to reduce the feature maps from 128 to 25. Finally, a 2D convolution is applied with 32 feature maps and 3x3 kernel size.

A down-sampling wide-activation block consists of two branches. One performs a simple max-pooling operation. The other contains 4 wide activation convolutional layers followed by a down-sampling 2D convolution with stride 2. The feature maps of each branch are summed.

An up-sampling wide-activation block also consists of two branches. The first consists of a single upsampling layer, which expands the input feature map depth from 32 to 128, applies a wide-activation convolution, and applies the depth-to-space upsampling operation. The other branch concatenates the input feature maps with those from encoder's skip-connection. These feature maps are then passed through 4 wide activation layers and an upsampling layer. The upsampled feature maps from both branches are then added together.

The final layer of the decoder branch is a 1x1 convolution operation that predicts the pixel values of the output image. In total, the full CNN model consists of 96 convolutional layers with 2.3 million trainable parameters.

Training and applying the CNN

To train such a large CNN on consumer GPUs, we used several tricks to reduce memory consumption. First, instead of using batch normalization to stabilize training, we used weight normalization. This enabled us to use smaller training batches (10 patches per batch) without training becoming slow or erratic. Second, we used gradient checkpointing to reduce GPU memory usage at the expense of longer computation time. We trained the network using the Adam stochastic optimizer for 100 epochs of 500 batches each. We initialize the learning rate at 1E-4 and decrement it by ¼ every 25 epochs.

Once trained, image paths and CTF parameters are read in from a STAR file. Each micrograph is then denoised, beginning with the same preprocessing pipeline used to generate training data. Instead of dividing the image into patches, the whole downsampled, phase-flipped image is passed through the trained CNN. Then, the image is padded with zeros in Fourier space and and a soft low-pass filter is applied at the Nyquist frequency of the down-sampled image. When inverse Fourier transformed, the final denoised image has the same pixel size and dimensions as the noisy input image and has the CTF corrected by phase-flipping.

Merging denoised and noisy images

Merging with noisy image and the denoised image is performed by high-pass filtering the noisy image and low-pass filtering the denoised image with complementary soft-edged Fourier

filters. Instead of Fourier padding into an array of zeros, the denoised image is effectively Fourier padded with the high-frequency information from the original noisy image. If not properly normalized, sharp changes in amplitude can occur in the band of spatial frequencies where the highpass and lowpass filters overlap (the merge band). To avoid this, the amplitudes of the denoised image are scaled by a normalization factor that keeps the standard deviation of amplitudes in the merge band the same after merging the denoised image into the noisy one. The frequency of the lowpass filter and the spectral width of the soft edge (both in angstroms) are parameters the user can specify.

MRC I/O operations use the library mrcfile. STAR file/CTF operations use our library pyem. All other image preprocessing is done using numpy, particularly the real FFT/IFFT routines. All CNN operations were performed with the library keras using the Tensorflow backend.

References

Abadi, Martín, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. 2016. "Tensorflow: A System for Large-Scale Machine Learning." In *12th Symposium on Operating Systems Design and Implementation,* 265–83.

Autzen, Henriette E., Alexander G. Myasnikov, Melody G. Campbell, Daniel Asarnow, David Julius, and Yifan Cheng. 2018. "Structure of the Human TRPM4 Ion Channel in a Lipid Nanodisc." *Science* 359 (6372): 228–32.

Batson, Joshua, and Loic Royer. 2019. "Noise2Self: Blind Denoising by Self-Supervision." *arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/1901.11365.

Baxter, William T., Robert A. Grassucci, Haixiao Gao, and Joachim Frank. 2009. "Determination of Signal-to-Noise Ratios and Spectral SNRs in Cryo-EM Low-Dose Imaging of Molecules." *Journal of Structural Biology* 166 (2): 126–32.

Bershad, N., and A. Rockmore. 1974. "On Estimating Signal-to-Noise Ratio Using the Sample Correlation Coefficient (Corresp.)." *IEEE Transactions on Information Theory / Professional Technical Group on Information Theory* 20 (1): 112–13.

Cheng, Yifan. 2015. "Single-Particle Cryo-EM at Crystallographic Resolution." *Cell* 161 (3): 450–57.

Cormier, Anthony, Melody G. Campbell, Saburo Ito, Shenping Wu, Jianlong Lou, James Marks, Jody L. Baron, Stephen L. Nishimura, and Yifan Cheng. 2018. "Cryo-EM Structure of the αvβ8 Integrin Reveals a Mechanism for Stabilizing Integrin Extension." *Nature Structural & Molecular Biology* 25 (8): 698–704.

Danev, Radostin, Dimitry Tegunov, and Wolfgang Baumeister. 2017. "Using the Volta Phase

Plate with Defocus for Cryo-EM Single Particle Analysis." *eLife* 6 (January). https://doi.org/10.7554/eLife.23006.

Frank, J., and L. Al-Ali. 1975. "Signal-to-Noise Ratio of Electron Micrographs Obtained by Cross Correlation." *Nature* 256 (5516): 376–79.

Glaeser, R. M. 1999. "Review: Electron Crystallography: Present Excitement, a Nod to the Past, Anticipating the Future." *Journal of Structural Biology* 128 (1): 3–14.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.

Grant, Timothy, and Nikolaus Grigorieff. 2015. "Measuring the Optimal Exposure for Single Particle Cryo-EM Using a 2.6 Å Reconstruction of Rotavirus VP6." *eLife* 4 (May): e06980.

Herzik, Mark A., Jr, Mengyu Wu, and Gabriel C. Lander. 2019. "High-Resolution Structure Determination of Sub-100 kDa Complexes Using Conventional Cryo-EM." *Nature Communications* 10 (1): 1032.

Jensen, G. J. 2001. "Alignment Error Envelopes for Single Particle Analysis." *Journal of Structural Biology* 133 (2-3): 143–55.

Kingma, Diederik P., and Jimmy Ba. 2014. "Adam: A Method for Stochastic Optimization." *arXiv [cs.LG]*. arXiv. http://arxiv.org/abs/1412.6980.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM*. https://doi.org/10.1145/306586.

Lehtinen, Jaakko, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. 2018. "Noise2Noise: Learning Image Restoration without Clean Data." *arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/1803.04189.

Ludtke, Steven J., and Wah Chiu. 2003. “Focal Pair Merging for Contrast Enhancement of Single Particles.” *Journal of Structural Biology* 144 (1-2): 73–78.

Odena, Augustus, Vincent Dumoulin, and Chris Olah. 2016. “Deconvolution and Checkerboard Artifacts.” *Distill*. https://doi.org/10.23915/distill.00003.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. 2015. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–41. Springer International Publishing.

Salimans, Tim, and Durk P. Kingma. 2016. “Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks.” In *Advances in Neural Information Processing Systems*, 901–9.

Schwartz, Osip, Jeremy J. Axelrod, Sara L. Campbell, Carter Turnbaugh, Robert M. Glaeser, and Holger Müller. 2019. “Laser Phase Plate for Transmission Electron Microscopy.” *Nature Methods* 16 (10): 1016–20.

Wong, W., and S. H. W. Scheres. 2019. “Cryo-EM Structure of the Plasmodium Falciparum 80S Ribosome Bound to the Anti-Protozoan Drug Emetine, Small Subunit.” https://doi.org/10.2210/pdb6okk/pdb.

Yu, Jiahui, Yuchen Fan, Jianchao Yang, Ning Xu, Zhaowen Wang, Xinchao Wang, and Thomas Huang. 2018. “Wide Activation for Efficient and Accurate Image Super-Resolution.” *arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/1808.08718.

Zhang, Kai. 2016. “Gctf: Real-Time CTF Determination and Correction.” *Journal of Structural Biology* 193 (1): 1–12.

Zhang, Kai, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017. “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising.” *IEEE*

Transactions on Image Processing: A Publication of the IEEE Signal Processing Society 26 (7): 3142–55.

Zheng, Shawn Q., Eugene Palovcak, Jean-Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard. 2017. "MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy." *Nature Methods* 14 (4): 331–32.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

DocuSigned by:
Palovcak, Eugene
20DA76539243463...
Author Signature

12/11/2019
Date